

# Trollfabriken und das Protestnetzwerk der russischen Opposition auf YouTube

kommunikation.medien

Open-Access-Journal  
für den wissenschaftlichen Nachwuchs

ISSN 2227-7277

Nr. 12 | 2020

<http://eplus.uni-salzburg.at/JKM>

DOI: 10.25598/JKM/2020-12.3



**Yulia Belinskaya**

## **Abstract**

*In diesem Artikel werden verschiedene Aspekte virtueller Öffentlichkeiten diskutiert, wobei der Schwerpunkt auf der Videoplattform YouTube liegt. Die Studie bezieht sich auf eine von dem russischen Oppositionellen Alexej Navalny auf YouTube geschaffene Community, die aufgrund ihrer hohen Wirkungskraft und Resonanz beispiellos ist. Im Januar 2020 besaß der Hauptkanal 3,2 Millionen Abonnentinnen und Abonnenten. Die Community dient als unabhängiges Forum, in dem nicht nur die Anhängerinnen und Anhänger, sondern auch die Kritikerinnen und Kritiker offen Dialog führen können. Mittels Kombination von digitalen Methoden werden die Dynamik und die Muster der Kommentare unter den auf dem YouTube-Kanal veröffentlichten Videos, die über die landesweiten Straßenproteste von 2017 bis 2019 in Russland berichten oder zur Teilnahme an ihnen aufrufen, analysiert. Es wurde versucht, die bezahlten Kommentatorinnen und Kommentatoren oder sogenannte Astroturfer auf Grundlage verschiedener Attribute (Aktivitätsniveaus des Kontos, Kommentaraktivität unter den Videos und Zeitleiste der Postings) zu identifizieren. Darüber hinaus wurde eine Stichprobe der potenziellen Astroturfer gezogen um deren strategische Rollen zu untersuchen, die sich wie folgt benennen lassen: Förderung, Unterbrechung, Respektlosigkeit und Misstrauen. Dieser Artikel versucht einen Grundstein für weitere Untersuchungen der virtuellen öffentlichen Sphäre auf YouTube zu legen.*

## **Keywords**

Virtuelle Öffentlichkeit, YouTube, Astroturfer, ideologisches Trolling, Trollfabriken, Identifizieren von Trollen

## 1. Einführung

Protestaktivitäten werden in Russland seit dem sogenannten Bolotnaja-Fall<sup>1</sup> immer häufiger gewaltsam unterdrückt. Im Jahr 2011 war das meistgenutzte soziale Medium der politischen Opposition *vk.com*<sup>2</sup> (vgl. Enikolopov et al. 2019). Drei Jahre später wurde dieses Medium von der Mail.ru-Gruppe übernommen – einem Unternehmen, das Oligarchen gehört, von denen angenommen wird, dass sie dem Kreml nahestehen (vgl. The Moscow Times 2014). Gleichzeitig begann Russland laut eines Berichts von Freedom House (2013) neben 22 anderen Ländern, auf bezahlte Kommentatorinnen und Kommentatoren in großem Umfang zu setzen. Ziel dieses Unterfangens ist es, die Meinung der Menschen online zu manipulieren. Dies stellt somit eine größere Gefahr für die Meinungsfreiheit und die weitere Entwicklung der Demokratie dar. Die Social-Media-Plattformen, die in der Wissenschaft als potenziell öffentliche Bereiche beschrieben werden, scheinen zu Instrumenten der sozialen Kontrolle geworden zu sein (vgl. Dahlberg 2007). Trotzdem gingen im Jahr 2017 zum ersten Mal seit fünf Jahren mehr als 100.000 Menschen in 83 russischen Städten auf die Straße, allesamt koordiniert mithilfe der auch als Protestnetzwerkmedium genutzten Plattform YouTube. Der russische Oppositionsaktivist und Gründer des „Fonds zur Korruptionsbekämpfung“<sup>3</sup> Alexej Navalny startete seinen ersten YouTube-Kanal im Juli 2013 nach dem Bolotnaja-Fall. Mit 3,2 Millionen Abonnentinnen und Abonnenten (Stand: Januar 2020) dient der Kanal hauptsächlich als Plattform zur Veröffentlichung der von der Agentur durchgeführten Antikorruptionsuntersuchungen, die von den traditionellen Medien nicht gezeigt werden. Die am meisten rezipierte Untersuchung, jene über den ehemaligen Präsidenten der Russischen Föderation Dmitri Medwedew, zählte mehr als 32 Millionen Aufrufe und 112.000 Kommentare.

Die von Alexej Navalny und seinem Team auf der YouTube-Plattform geschaffene Community ist aus den folgenden vier Gründen beispiellos: erstens wegen der extrem hohen Wirkungskraft und Resonanz des YouTube-Kanals einer politischen Figur, die in keinen staatlichen Medien auftreten darf (vgl. Glazunova 2020); zweitens wegen des Formats des Materials, das in Form von Dokumentarfilmen präsentiert wird; drittens wegen des Raumes, der dabei zur Verfügung gestellt wird, um die Korruptions-Untersuchungen zu kommentieren und zu kritisieren. Das

---

<sup>1</sup> Dabei handelt es sich um den vom Untersuchungsausschuss der Russischen Föderation untersuchten strafrechtlichen Fall von Massenunruhen und Gewalt gegen Polizeikräfte, der während der Protestserie Ende 2011/Anfang 2012 als Reaktion auf den Wahlbetrug 2011 stattfand. Die Anklagen gegen die Teilnehmerinnen und Teilnehmer wurden als „politisch motiviert“ eingeordnet und unter der Bezeichnung „Bolotnaja-Fall“ bekannt (vgl. Bodnar et al. 2013).

<sup>2</sup> vk.com ist ein mehrsprachiges soziales Netzwerk, das im Jahr 2006 entwickelt wurde und eine Alternative zu Facebook darstellt. Hauptsächlich ist es in Russland und den Staaten der GUS weit verbreitet.

<sup>3</sup> Der „Fonds zur Korruptionsbekämpfung“ ist eine russische nichtstaatliche gemeinnützige Organisation, die korruptionsbezogene Straftaten in den höchsten Behörden untersucht und veröffentlicht.

Motto des Kanals lautet „Hier wird die Wahrheit gesprochen“, was zu einer uneingeschränkten Teilnahme und offenen Diskussion einzuladen scheint. Viertens wurde die Plattform parallel zu den Facebook-Gruppen, die die lokalen Protestorganisationen anführten, zum wichtigsten Protest-Organisationsnetzwerk Russlands.

Ziel dieses Artikels ist es, mittels einer Kombination verschiedener digitaler Methoden die Dynamik und die Muster der Kommentare unter den im YouTube-Kanal von Alexej Navalny veröffentlichten Videos zu analysieren. Dabei sollen die bezahlten Kommentatorinnen und Kommentatoren und ihre Rolle in der Diskussion anhand des beobachteten Online-Verhaltens identifiziert werden. Folgende Forschungsfragen leiten die einzelnen Schritte dieser Studie: Wie reagiert das Publikum auf Videos, die versuchen, ein Protestnetzwerk zu schaffen? Welche dominanten Themen charakterisieren diese Reaktionen? Welche Strategien des *ideologischen Trollings* können im Verhalten potenziell bezahlter Trolle identifiziert werden? Welche Kriterien können bei der Identifizierung von koordiniertem Troll-Verhalten auf YouTube nützlich sein?

## 2. Theorie und Forschungsstand

In den folgenden Abschnitten wird das Konzept der Öffentlichkeit, seine Verwendung im digitalen Bereich und die Möglichkeit der Operationalisierung des Konzepts hinsichtlich spezifischer Social-Media-Plattformen wie YouTube diskutiert. Darüber hinaus werden das Phänomen der bezahlten Trolle – oder *Astroturner* – sowie einige frühere Versuche, eine Methode zu entwickeln, um diese in sozialen Medien zu entdecken, vorgestellt. Diese Themen bilden den theoretischen und methodischen Rahmen für das Forschungsprojekt.

### 2.1 Konzept der Öffentlichkeit und Kritik

Nach Habermas (1989: 49) ist die Öffentlichkeit „ein Bereich unseres gesellschaftlichen Lebens, in dem sich etwas, das sich der öffentlichen Meinung annähert, bilden kann. Der Zugang ist allen Bürgern garantiert“. Dieser Bereich funktioniert nicht als legale staatliche Institution, sondern ist ein wesentlicher Grund für eine funktionierende Demokratie. Die öffentliche Meinung wird durch rationale Debatten und logische Diskussionen gebildet (vgl. Habermas 1981). Die Zeitungen, das Fernsehen und andere Medien fungieren als Manifestation einer öffentlichen Sphäre, in der Standpunkte vertreten und diskutiert werden können. Der Kapitalismus und die ungleiche Verteilung des Reichtums haben für die allgemeine Öffentlichkeit jedoch Barrieren für den Eintritt in die öffentliche Sphäre geschaffen. Garnham (1990) weist auf die Unabhängigkeit solcher Sphären von Wirtschaft und Staat hin. Im Falle von Massenmedien,

die ihre kommerzielle und politische Funktion gleichzeitig ausüben, wird diese Unabhängigkeit problematisch: „Foren zur Diskussion politischer Standpunkte in Film und Fernsehen stehen unter der Kontrolle professioneller Medienorganisationen und ihre Formate schränken den öffentlichen Input häufig ein.“ (Milliken et al. 2008: 1)<sup>4</sup>

Verschiedene Wissenschaftlerinnen und Wissenschaftler haben die Problematik der öffentlichen Sphäre in der modernen Gesellschaft erkannt. Sie wurde als gefährdet (vgl. Johnson 2006) oder in multikulturellen Gesellschaften sogar als abwesend (vgl. Walzer 1999) beschrieben. Als Begründung hierfür wird angeführt, dass es keine Möglichkeit gibt, sich zusammen auf allgemeine Probleme und Lösungen zu einigen, die für die Gesellschaft als Ganzes vorteilhaft sind. Das Konzept selbst wurde wegen seines Klassenunbewusstseins sowie der Fokussierung auf „die bürgerliche Öffentlichkeit als Idealtypus“ (Garnham 2007: 207) aktiv kritisiert. Darüber hinaus wurden auch Fragen nach *race* (Jacobs 1999) und Geschlecht (Fraser 1990) in Bezug auf die *inclusivity* der bürgerlichen Öffentlichkeit, die Minderheiten eher davon abhält, ihre Anliegen zu artikulieren, aufgeworfen.

## 2.2 Virtuelle Öffentlichkeit

Das Potenzial des Internets, zu einer neuen, für alle zugänglichen Öffentlichkeit zu werden, die in der Lage ist, „unverfälschte“ Kommunikation zu ermöglichen, wird schon seit Längerem in der Literatur diskutiert (vgl. unter anderem Holmes 1997; Papacharissi 2002; Dahlberg 2007; Geiger 2009). So stellt die Internet-Technologie den Bürgerinnen und Bürgern ein Werkzeug zur sozialen Interaktion mit dem Potenzial einer globalen Reichweite zur Verfügung (vgl. Bolognani 2011). Auch wenn nach wie vor eine „digitale Kluft“ (NTIA 1995) besteht, hat das Internet eine sehr „niedrige Zugangsbarriere“ (Edgerly et al. 2009: 5). Zusammen mit der Möglichkeit, Inhalte hochzuladen, zu kommentieren und zu kritisieren, stellt dies ein weiteres Merkmal von Öffentlichkeit dar und betrifft beispielsweise *inclusivity* und Gleichheit. Dahlberg (2001) zufolge könnte das Internet die Hauptmerkmale von Öffentlichkeit erfüllen: Autonomie vom Staat und wirtschaftliche Unabhängigkeit; Meinungsaustausch und offene Kritik; Reflexivität, die sich in der Suche nach einem besseren Argument widerspiegelt; Rollenübernahme, die respektvolles Zuhören einschließt; Aufrichtigkeit; *inclusivity* und Gleichheit. Die ersten beiden Merkmale zeigen die Deliberation und die anderen vier verweisen auf die Qualität dieser.

Die häufig an virtuelle Öffentlichkeiten geknüpfte Hoffnung, einen sozialen Wandel zu fördern und zur Deliberation zu ermutigen, ist in der akademischen Diskussion eines der am meisten

---

<sup>4</sup> Alle Zitate wurden von der Autorin aus dem Englischen ins Deutsche übersetzt.

erörterten Themen (vgl. Papacharissi 2008; Bolognani 2011). Das demokratische Potenzial des Internets, das global funktioniert und weitgehend frei von Regulierungen ist, wurde als utopische Wiedergeburt des habermasianischen Konzepts erachtet (Milliken et al. 2008). Die Skeptikerinnen und Skeptiker verwiesen auch auf das kommerzielle Potenzial der Internetplattformen, das neue Barrieren und Hemmnisse zur Folge hätte (vgl. Sparks 2001). Thomas und Wyatt (1999) sagten etwa voraus, dass die großen Konzerne einen Weg zur Kommerzialisierung des Cyberspace finden würden, der jedoch gleichzeitig das Tor für die globalen, kulturellen und politischen Interaktionen öffnen würde. Mit dem Aufstieg globaler Konzerne wie Facebook und Google ist die wirtschaftliche Autonomie der virtuellen Öffentlichkeit von einigen Forscherinnen und Forschern neu infrage gestellt worden (vgl. Valtysson 2012).

Ein weiteres wichtiges Thema in der Diskussion um die virtuelle Öffentlichkeit ist die Gefahr der Fragmentierung. Habermas (2006) argumentiert in diesem Zusammenhang folgendermaßen: „Die computervermittelte Kommunikation hat in der Öffentlichkeit kaum mehr als eine parasitäre Rolle zu spielen, was weitgehend auf die Art und Weise zurückzuführen ist, wie internetbasierte Diskursgemeinschaften die Öffentlichkeit fragmentiert haben.“ Zudem ist das Internet nicht frei von Politik. „Der Aufstieg des Internets erweitert den Raum für demokratische Beteiligung sowie Debatten und schafft neue öffentliche Räume für politische Interventionen.“ (Kellner 2000: 13) Aufgrund des schwindenden Vertrauens in die traditionellen Medien (Freedom House 2013) erscheint das habermasianische Konzept der Öffentlichkeit wie ein unerreichbarer utopischer Zustand. Die Debatte über die Möglichkeiten des Internets, die Eintrittsbarrieren zu überwinden, wird seit dem Aufkommen der Technologie intensiv geführt. Die Fragmentierung des Publikums, die überwältigende Menge an Informationen und die Existenz von Echokammern (Colleoni et al. 2014) stellen dabei die konzeptionellen Unklarheiten dar. Trotzdem sollte das Potenzial einer virtuellen Öffentlichkeit für die Zivilgesellschaft nicht unterschätzt werden. Die Möglichkeit der Anwendung des Konzepts der Öffentlichkeit auf autoritäre Demokratien ist ebenfalls Gegenstand dieses Artikels.

Angelehnt an Habermas (1981) wird YouTube hier als kommunikativer Raum betrachtet, innerhalb dessen die Teilnahme an Diskussionen unbegrenzt ist. Habermas hebt den Wert der öffentlichen Meinung hervor und fordert die Schaffung von Räumen, in denen die Rede frei von „zwanghaften Verzerrungen“ ist und „offen kritisiert werden kann“ (Godin et al. 2007: 454). Man kann argumentieren, dass es im Fall von YouTube gewisse Einschränkungen gibt, die von der Plattform selbst auferlegt werden – zum Beispiel sind anonyme Kommentare nicht erlaubt. Es gibt jedoch unzählige, nicht auffindbare pseudoanonyme Nutzerinnen und Nutzer und nur ein minimales Gatekeeping.

### 2.3 YouTube als Öffentlichkeit

2019 wurde YouTube zur meistbesuchten Website in den USA und zur am zweitmeisten genutzten Website der Welt und ließ damit Facebook, Twitter und Amazon hinter sich. Durch die einfachen Zugangsmöglichkeiten und das Erreichen eines globalen Publikums „ist YouTube zu einem politischen Forum geworden, in dem [...] Bürger Inhalte produzieren, die von den Mainstream-Medien aufgegriffen werden“ (Edgerly et al. 2009: 2). Milliken et al. (2008) fordern eine neue Konzeptualisierung der öffentlichen Sphäre angesichts der Instrumente, die von Plattformen wie YouTube zur Verfügung gestellt werden, und der Art und Weise, wie Menschen dadurch ihre politischen Ansichten ausdrücken können. Anthony und Thomas (2010) reflektierten über die Idee der auf Nutzerinhalten basierten Plattformen und betonen, dass die Möglichkeit, die Interessen verschiedener Nutzerinnen- und Nutzergruppen zu repräsentieren, den elitären Charakter der öffentlichen Sphäre überwindet.

Die Anonymität auf einer Plattform wie YouTube eröffnet jedoch die Möglichkeit für unwahre Aussagen und Manipulationen. Das bedeutet, dass dieser virtuelle Kommunikationsraum auch von undemokratischen Regierungen genutzt werden kann, um die öffentliche Meinung jenseits der staatlichen Kanäle zu steuern. Diese Regierungen beeinflussen den Diskurs und verschieben die Tagesordnung, indem sie bezahlte Online-Kommentatorinnen und -Kommentatoren oder, wie sie in Russland üblicherweise genannt werden, „Kremlinbots“, „Web-Brigaden“ oder „Trollfabriken“ einsetzen (vgl. Aro 2016). Dieses Phänomen wurde am Beispiel der chinesischen „50-Cent-Partei“ bereits von verschiedenen Wissenschaftlerinnen und Wissenschaftlern untersucht (Han 2015; King et al. 2017).

### 2.4 Trollfabriken

In einem speziellen Beschluss warf das Europäische Parlament die Frage der bezahlten Internet-Trolle auf, die von der russischen Regierung eingesetzt werden, „um demokratische Werte in Frage zu stellen, Europa zu spalten, innere Unterstützung zu sammeln und die Wahrnehmung gescheiterter Staaten in der östlichen Nachbarschaft der EU zu schaffen“ (European Parliament 2016). Bradshaw und Howard (2017: 4) erklären: „Viele Regierungen wenden heute erhebliche Ressourcen auf und beschäftigen eine große Zahl von Menschen, um Inhalte zu generieren, Meinungen zu steuern und mit dem ausländischen und inländischen Publikum in Kontakt zu treten.“ Das Hauptziel solcher Aktivitäten ist es, die Agenda zu verschieben und die öffentliche Reaktion auf bestimmte Themen zu lenken. Dazu gehören die Aktivitäten der 50-Cent-Truppen in China, deren Aufgabe „das Veröffentlichen von ablenkenden Kommentaren in den hochrangigen Oppositionsblogs (Ananyev & Sobolev 2017) sowie das Zurückschreiben

und Umsetzen von regierungsfreundlichen Botschaften“ (Sanovich 2017: 9) ist. Zhang et al. (2013) definieren diesen Prozess als „Astroturfing“, was bedeutet, bestimmte Ansichten im Namen einer Organisation zu kommunizieren. Die Kommentatorinnen und Kommentatoren werden mit dem strategischen Ziel eingestellt, Desinformationen zu verbreiten, sodass die kundengegebenen Ansichten von einem breiteren Publikum unterstützt werden (vgl. Miller 2016). Wie Aro (2016: 122) betont, „haben aggressive pro-russische Trollkampagnen die öffentliche Debatte manipuliert und die Bürger zum Schweigen gebracht“. In der russischen Sprache wird das Wort „Troll“ synonym für den Begriff *Astroturfer* verwendet. Hier muss herausgestellt werden, dass Trolle als Accounts verstanden werden, die von Menschen geführt werden – im Gegensatz zu Bot-Konten, die computergesteuert sind (vgl. Linvill & Warren 2019).<sup>5</sup>

Wie von Zhdanova und Orlova (2017) beschrieben, wurde im September 2013 die russische Enthüllungsjournalistin Alexandra Garmazhapova von der russischen Internetforschungsagentur mit Sitz in Olgino bei St. Petersburg angestellt. In dem von der *Nowaja Gaset*a veröffentlichten Material wird diese Agentur als „Trollfabrik“ bezeichnet. „Kommentare ohne Links werden auch direkt in Foren gepostet und jeder Troll muss während einer 12-Stunden-Schicht Hunderte von Kommentaren produzieren.“ (Aro 2016: 125) Der Kreml streitet allerdings jede Verbindung zu den Trollfabriken ab (vgl. Kirby 2018). In einem Interview beschrieb ein ehemaliger Mitarbeiter die radikale Veränderung der Anforderungen in Trollfabriken (vgl. Volchek 2018). Zuvor war es wichtig, die Sachinformationen bereitzustellen und Kommentare zu veröffentlichen, „die scheinbar genaue Listen von Quellen enthalten, die die Glaubwürdigkeit des Textes unterstreichen“ (Aro 2016: 125). Die neue Tagesordnung soll jedoch zu einer übermäßigen Menge an Kommentaren führen, die starke negative Konnotationen erzeugen und insbesondere mit den Oppositionellen in Zusammenhang gebracht werden (vgl. Volchek 2018).

## 2.5 Identifizierung von Trollen

Die Erkennung von Bots setzt die Anwendung von automatisierten Computerprogrammen voraus – in der Regel überwachte oder unbeaufsichtigte Lernalgorithmen (Stukal et al. 2017). Um zu unterscheiden, ob ein Mensch oder ein Bot einen Twitter-Account kontrolliert, verwenden Badawy et al. (2018) die folgenden Kriterien: (1) übermäßige Aktivität; (2) keine Anpassung des Kontos; (3) keine GPS-Metadaten vorhanden; (4) zahlenmäßiges Verhältnis von Retweets zu Original-Tweets; (5) zahlenmäßiges Verhältnis von Followern zu Followees; (6)

---

<sup>5</sup> Anmerkung: Bot-Konten werden im weiteren Verlauf als „Bots“ bezeichnet.

Erstellungsdatum des Kontos; (7) Zufälligkeit des Benutzernamens. Mit einer gewissen Vorsicht könnten diese Kriterien auch auf die Erkennung von Trollen angewandt werden. Nach Han (2015) sind Trolle wenig motiviert, ihre Arbeit gut zu machen, deshalb kopieren sie die Posts von anderen, um ihre Arbeitsbelastung zu reduzieren. Darüber hinaus schreiben *Astroturfer* oft politische Kommentare unter verschiedenen Spitznamen (Miller 2016), wodurch die Kriterien (2) bis (7) bei ihrer Erkennung hilfreich sein können.

Im Allgemeinen erweist sich das Erkennen von Trollen oder *Astroturfern* aber als eine schwierige Aufgabe für das maschinelle Lernen. Miller (2016) erläutert diesbezüglich: „Astroturfer sind, anders als Bots, Menschen, die üblicherweise manuell posten und oft kundenspezifische Nachrichten erstellen.“ (ebd.: 2) Sie verbergen nicht bloß ihre Identität und posten Kommentare anonym, sondern versuchen vielmehr, als gewöhnliche Benutzerinnen und Benutzer zu erscheinen (vgl. Peng et al. 2016). Die Aufgabe besteht somit nicht nur darin, verdächtige Aktivitäten zu erkennen, sondern „eine Klasse von Menschen von einer anderen zu unterscheiden“ (Miller 2016: 2). Mihaylov et al. (2015) untersuchten unter anderem die Zeitachse, basierend auf der Annahme, dass Trolle einer bezahlten Vollzeitätigkeit nachgehen und zu den Ersten gehören würden, die sich zu einem Thema äußern. Um einen Weg zu finden, die bezahlten Trolle von jenen Menschen zu unterscheiden, deren Meinung tatsächlich mit der Regierungspropaganda übereinstimmt, verwenden die Wissenschaftlerinnen und Wissenschaftler *Ground-Truth*-Daten, das heißt durchgesickerte Dokumente wie Anweisungen für *Astroturfer*, Listen von Trollen, IP-Adressen in Kombination mit verschiedenen Techniken wie der Zuweisung der Autorinnen- und Autorenschaft (vgl. Peng et al. 2016) oder eine Sentimentanalyse (vgl. Seah et al. 2015). Mihaylov et al. (2015: 310) schlagen für die Identifizierung der bezahlten Trolle eine weitere Möglichkeit vor: „Wir gehen davon aus, dass ein Nutzer, der von mehreren Personen als Troll bezeichnet wird, wahrscheinlich auch einer ist.“ Auch Tomaiuolo et al. (2020) betonen die Bedeutung der Reaktionen der Gemeinschaft: Als Hauptakteurinnen und Hauptakteure der öffentlichen Sphäre erstellen und verbreiten die Nutzerinnen und Nutzer nicht nur Inhalte, sondern tragen auch die Verantwortung, die Debatte auf der Plattform zu regulieren.

Sobolev (2019) zufolge verfolgen staatlich geförderte Kommentatorinnen und Kommentatoren im Allgemeinen folgende drei Ziele: eine Anti-Oppositions- oder Pro-Regierungs-Diskussion zu unterstützen, die Nutzerinnen und Nutzer von gefährlichen Themen abzulenken oder die Diskussion vollständig zu unterbrechen, indem sie die Nutzerinnen und Nutzer mit einigen radikalen Aussagen provozieren und Verbote auslösen (Beförderung, Ablenkung und Handlungsaufforderung). Im Zusammenhang mit Protesten kann dies in Form von Aufrufen zu extremistischen Aktionen oder zur Radikalisierung einer friedlichen Versammlung geschehen.

Ablenkung ist eine Aktivität, die darauf abzielt, die Aufmerksamkeit von der Kritik abzulenken, zum Beispiel mit der Technik des „Whataboutismus“, bei den Fragen wie „und was ist mit den USA?“ gestellt werden. Hardaker (2013) argumentiert, dass Trolle vier miteinander verbundene Strategien anwenden: Täuschung, Aggression, Unterbrechung und Erfolg (im Mittelpunkt der Aufmerksamkeit zu stehen). Zelenkauskaitte und Niezgodna (2017) unterscheiden drei Hauptmodi des Trollens: Unterbrechung, Misstrauen und Respektlosigkeit. King et al. (2017) differenzieren zwischen Cheerleading, argumentativem und nicht-argumentativem Lob oder Vorschlägen, sachlicher Berichterstattung und der Verspottung anderer Länder.

### 3. Methode

Laut Hussain et al. (2018) besteht ein Mangel an Forschung hinsichtlich Video-Sharing-Plattformen wie YouTube, während in Bezug auf Twitter und Facebook viel Forschungsarbeit geleistet wird. Es gibt jedoch keine gängige Methode, die sich auf die Untersuchung von Kommentaren auf YouTube anwenden ließe. In diesem Beitrag wird versucht, eine methodische Vorgehensweise zu entwickeln, die dazu beitragen könnte, bezahlte Internet-Trolle aufzuspüren. Es soll zudem möglich sein, diese Vorgehensweise auf den YouTube-Kommentarabschnitt anzuwenden. Die Methoden zur Analyse von Microblogging-Plattformen wie Twitter oder von anderen sozialen Netzwerken wie Facebook, die hauptsächlich für private oder geschäftliche Zwecke genutzt werden (vgl. Logan 2014), konnten nur begrenzt angewandt werden, weil sich YouTube als soziales Netzwerk deutlich von solchen Plattformen unterscheidet. Die Microblogging-Plattformen sind von den sozialen Offline-Verbindungen zwischen den Nutzerinnen und Nutzern geprägt, da der Feed sich zumeist aus Personen zusammensetzt, die die Nutzerinnen und Nutzer im realen Leben kennen (vgl. Georgakopoulou & Spilioti 2015). Bei YouTube hingegen besteht nahezu vollständige Anonymität: „Um Kommentare hinzufügen zu können, müssen sich die Nutzer registrieren und ihren Spitznamen, der in der Regel nicht preisgegeben wird, sowie grundlegende Informationen [...] angeben, die jedoch erfunden sein können“ (Dyrel 2012: 35).

#### 3.1 Stichprobe

Eine Stichprobe von 1.131 Textkommentaren (aus einem Pool von 63.384 Kommentaren), die unter drei verschiedenen Videos veröffentlicht wurden, bildet den Gegenstand dieser Forschung. Zunächst wurden nur jene Videos ausgewählt, die über die folgenden drei Proteste berichteten und zu deren Organisation beitrugen:

- Der Protest am 26.03.2017 „Er ist für uns kein Zar“ richtete sich gegen die Korruption, die in der Untersuchung über den Premierminister Dmitri Medwedew aufgedeckt wurde.
- Der Protest am 09.09.2018 mobilisierte gegen die Reform des Rentensystems.
- Der Protest am 27.07.2019 richtete sich gegen die Beschränkungen, die den Kandidatinnen und Kandidaten für den Moskauer Stadtrat auferlegt wurden.

Die Auswahl der willkürlichen Stichproben bewirkte, dass die Videos zeitlich auseinanderliegen und jeweils ein Jahr (2017, 2018, 2019) repräsentieren. Dadurch ist es möglich, Vergleiche zu ziehen und einige etablierte Trends zu diskutieren. Darüber hinaus haben jene Videos, die zum Protest aufrufen, weniger Klicks, jedoch mehr Kommentare, als die Videos, die Untersuchungen thematisieren. Von jeder dieser drei Gruppen (jeweils mehr als 30 Videos) wurden nach dem Zufallsprinzip drei Videos ausgewählt, um eine repräsentative Stichprobe zu erhalten:

- Video 1, veröffentlicht am 23. März 2017
- Video 2, veröffentlicht am 30. August 2018
- Video 3, veröffentlicht am 25. Juli 2019

Die Kommentare wurden über das offene Datenerhebungsinstrument<sup>6</sup> gesammelt (Rieder 2015). Zu den Daten gehören die Benutzer-ID, der Benutzername, der Zeitpunkt, zu dem eine Nachricht gepostet wurde, der Kommentar, die Anzahl der Antworten sowie die Benutzer-ID und ein Benutzername, an den dieser Kommentar gerichtet ist. Die Zufallsstichproben wurden mit 95 % Konfidenzintervall extrahiert: Video 1 – 13.343 Kommentare / Stichprobengröße 374; Video 2 – 27.371 / 379; Video 3 – 22.670 / 378.

### 3.2 Thematische Analyse

Die thematische Analyse diente als wichtiger erster Schritt der Forschung, da sie dazu beitrug, die Verhaltensmuster des Publikums und seine Einstellungen zu den Videos aufzudecken. Mit diesen Mustern war es möglich, die Eigenheiten der virtuellen Öffentlichkeit aufzuzeigen und zu diskutieren. Sobolev (2019) zufolge besteht eine der Hauptaufgaben der Trolle darin, eine große Menge an Kommentaren zu erstellen, die Anti-Oppositions- oder Pro-Regierungs-Diskurse schaffen bzw. zu diesen beitragen. Um die Trolle in diesen Diskursen zu platzieren, war es daher unerlässlich, diese zuerst zu kartieren.

---

<sup>6</sup> YouTube Data Tools (<https://tools.digitalmethods.net/netvizz/youtube/>).

Die Kommentare wurden einer thematischen Analyse unterzogen, die auf dem Grounded Theory-Ansatz basiert (vgl. Strauss & Corbin 1994). Zunächst wurde der Inhalt der Kommentare unter dem ersten Video uneingeschränkt offen kodiert und konzeptualisiert, um auftauchende Themen herauszuarbeiten. Die nachfolgende Forschungsfrage leitete diese Phase der Forschung:

*FF1: Wie reagiert das Publikum auf die Videos, die versuchten, ein Protestnetzwerk zu schaffen? Welche dominanten Themen charakterisieren diese Reaktionen?*

Die identifizierten Themen wurden in sechs Hauptkategorien gruppiert, innerhalb derer mehrere Unterkategorien identifiziert wurden. Zu den vier größeren Kategorien zählen:

- Unterstützung von Navalny oder des Protestes;
- negative Kommentare gegenüber Navalny, den Protestierenden oder dem Protest;
- negative Kommentare, die auf Putin, Medwedew, die Regierung und das Regime abzielen;
- Unterstützung von ebendiesen.

Als fünfte Kategorie wurde „Diskussion“ aufgenommen, die Diskussionen und Fragen zum Thema der Videos umfasst – etwa in Bezug auf die Rentenreform und die Moskauer Stadtratswahlen. Die sechste Kategorie mit dem Titel „andere oder irrelevant“ enthält verschiedene Reaktionen wie Emoticons, Pluspunkte, Diskussionen über andere Themen, Beleidigungen gegenüber bestimmten Benutzerinnen und Benutzern, Beschwerden über YouTube-Algorithmen oder -Regeln, Witze und andere Kommentare, die nicht als unterstützend oder negativ eingestuft werden konnten.

### 3.3 Troll-Probenentnahme

Da für diese Forschung keine *Ground-Truth*-Daten zur Verfügung standen, wurde die von Mihaylov et al. (2015) verwendete Methode verfolgt und alle Nutzerinnen und Nutzer identifiziert, die von anderen Mitgliedern der Gemeinschaft als *Astroturfer* bezeichnet wurden. Dies diente zur Extrahierung einer Stichprobe potenziell<sup>7</sup> bezahlter Trolle. Die allgemeine Stichprobe von 63.384 Kommentaren wurde auf die Verwendung der folgenden Wörter hin durchsucht: „Kremlinbot“, „Bot“, „Troll“, „Trollfabrik“, „Savushkina“ (der Name der Straße, in der sich die Trollfabrik vermutlich befindet) oder „Olginez“ (der Name des Dorfes). Die Ergebnisse

---

<sup>7</sup> Bei dieser Forschung war es nicht möglich, die Trolle mit 100%iger Sicherheit zu identifizieren. Daher hat die Autorin bei deren Beschreibung und beim Identifikationsprozess immer die *potenziellen* Trolle im Blick.

wurden manuell auf ihren Kontext hin überprüft, etwa wenn das Wort „Troll“ oder „Bot“ verwendet wurde, um die sich wiederholenden Kommentare von computergesteuerten Konten zu identifizieren. Die Liste mit 193 Benutzerinnen- und Benutzernamen wurde in drei Videos identifiziert. Es ist anzumerken, dass es bei dieser Extraktionstechnik gewisse Einschränkungen gibt. Da die Empfängerin bzw. der Empfänger nur auf der ersten Diskussionsebene sichtbar sein konnte, waren etwa 40 % der Beschuldigungen der Trolle nicht zu identifizieren. Weitere 7 % der Benutzerinnen und Benutzer konnten nicht nachverfolgt werden, da ihre Konten gelöscht wurden. Daher umfasste die endgültige Stichprobe 127 Benutzernamen. Die nachfolgende Forschungsfrage leitete die zweite Phase der Studie:

*FF2: Welche Strategien des ideologischen Trollings können im Verhalten potenziell bezahlter Trolle identifiziert werden?*

Um diese Forschungsfrage zu beantworten, wurde eine thematische Analyse auf der Grundlage der verschiedenen im Kapitel „Identifizierung von Trollen“ diskutierten Arbeiten durchgeführt und die verschiedenen Modi zusammengeführt und angepasst. So gelang es, ein Codebuch zu erstellen, das dabei helfen sollte, die von den bezahlten Trollen angewandten Strategien zu erkennen. Die Kodierung wurde durch die vorhandene Literatur vorangetrieben: Nachdem verschiedene Strategien identifiziert wurden, die von den Autorinnen und Autoren bei der Trainingsstichprobe von 30 Kommentaren beschrieben wurden, war es möglich, die Kategorien anzupassen. Das daraus resultierende Codebuch besteht nunmehr aus vier Hauptmodi bzw. Strategien: Förderung, Unterbrechung, Respektlosigkeit und Misstrauen.

### **3.4 Weitere Untersuchung**

*FF3: Welche Kriterien können bei der Identifizierung von koordiniertem Trollverhalten auf YouTube nützlich sein?*

Da keine *Ground-Truth*-Daten verfügbar sind, kann die Aufgabe der Erkennung von bezahlten Trollen möglicherweise nicht mit Sicherheit erfüllt werden. Daher hat die Studie explorativen Charakter und dient dem Zweck der Unterscheidung und der Erkennung der Muster von koordiniertem und verdächtigem Verhalten sowie dessen Kategorisierung. Die Stichprobe von 127 Konten wurde untersucht, um folgende Daten zu ermitteln: das Erstellungsdatum des Kontos, das Aktivitätsniveau auf dem Kanal der Nutzerin oder des Nutzers (einschließlich des Grades der Anpassung, der Anzahl der Abonnements und der Frage, ob Videos auf die Plattform hochgeladen wurden). Die Konten wurden anschließend binär als aktiv oder inaktiv kodiert. Das Konto wurde als inaktiv eingestuft, wenn es kein oder nur ein Abonnement und keine oder nur eine grundlegende Anpassung gab sowie keinerlei Videos hochgeladen wurden.

Anschließend wurden die Aktivitäten der Benutzerinnen und Benutzer unter den Videos und die Zeitleiste ihrer Postings manuell überprüft. Wenn die Benutzerin oder der Benutzer 5 bis 15 Kommentare zu einem bestimmten Video abgegeben hatte, wurde die Aktivität des Postings als mäßig und bei mehr als 15 als exzessiv eingestuft. Darüber hinaus wurde die Zeitachse der Beiträge untersucht. Laut Tomaiuolo et al. (2020) ist die Zeit ein wichtiger Indikator, insbesondere wenn die Kommentare täglich gepostet werden. Dies wird auch von Aro (2016) unterstrichen und in Interviews mit ehemaligen Arbeiterinnen und Arbeitern von Trollfabriken bestätigt. Wenn die Nutzerinnen und Nutzer durchgehend an mindestens vier aufeinanderfolgenden Tagen posteten, wurden sie als „Tageskommentatorinnen bzw. -kommentatoren“ bezeichnet, wenn die Kommentare in den ersten fünf Stunden nach der Veröffentlichung des Videos erschienen, wurden sie zudem mit „unter den ersten Kommentatorinnen bzw. Kommentatoren“ gekennzeichnet.<sup>8</sup>

Unter Berücksichtigung der einzigen für diesen Fall verfügbaren *Ground-Truth*-Daten – Interviews, die von den ehemaligen Arbeiterinnen und Arbeitern der Trollfabriken gegeben wurden, und theoretische Beschreibungen von Han (2015) und Miller (2020) – sind die Aktivität des Senders und der Umfang der von dem jeweiligen Konto durchgeführten Aktivitäten wichtige Indikatoren, die in Kombination mit anderen Kriterien nützlich sein könnten.

## 4. Ergebnisse

### 4.1 Reaktionen auf die Videos

Mittels einer thematischen Analyse wurden die Dynamiken und die Muster der Kommentare unter den Videos, die auf dem YouTube-Kanal von Alexej Navalny veröffentlicht wurden, untersucht. Die negativen oder missbilligenden Reaktionen gegenüber Navalny machen bei den Videos 1 bis 3 24 %, 6 % und 7 % aus (siehe Abbildung 1). Dabei haben sich folgende Themen herauskristallisiert:

---

<sup>8</sup> Zusätzlich wurden die visuellen Netzwerke mithilfe der Gephi-Software aufgebaut. Die Netzwerkanalyse kann weiterverwendet werden, um die Verbindungen zwischen den Benutzerinnen und Benutzern zu untersuchen und die ‚verdächtigen‘ Konten aufzuzeigen. Die Netzwerkanalyse ist jedoch nicht Bestandteil der vorliegenden Arbeit und soll für zukünftige Studien zur Verfügung stehen.

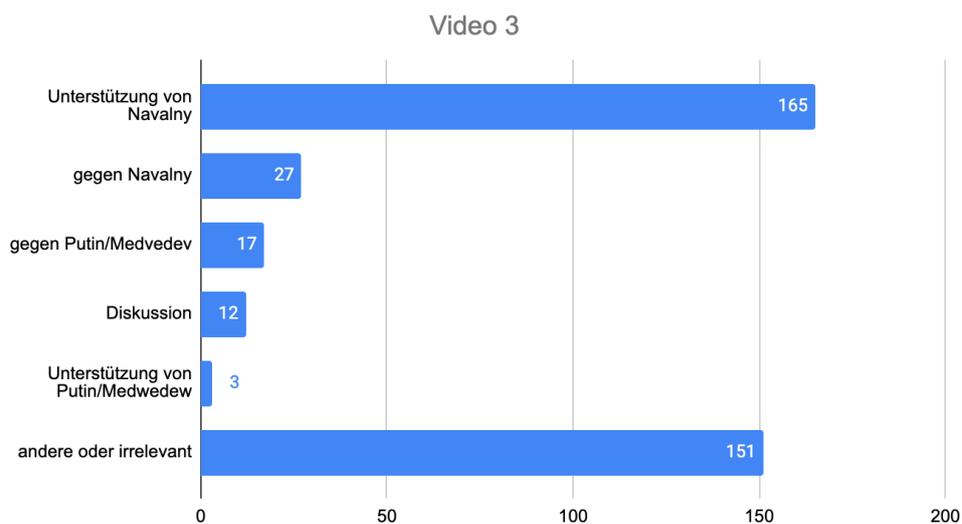
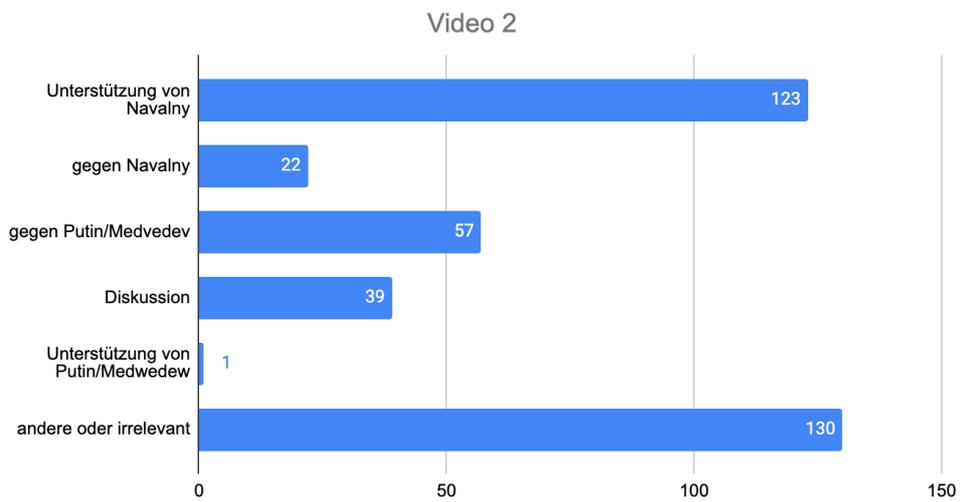
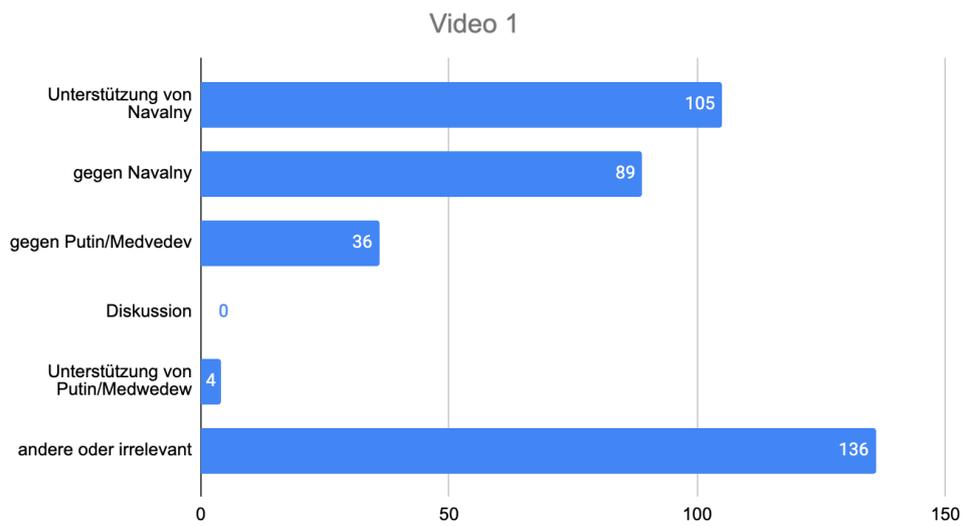


Abbildung 1: Hauptgruppen von Antworten auf die Videos 1 bis 3 (eigene Darstellung)

(1) Navalny bringt die Unterstützerinnen und Unterstützer in Gefahr und besucht die Proteste nicht selbst

„Ja, in der Tat, wenn irgendeine Art von Bewegung stattfindet, lässt Navalny seine Unterstützer fallen.“ In den wenigen Fällen, in denen Navalny am Tag der Proteste oder einige Zeit zuvor wegen der Organisation der Proteste verhaftet wurde, begann er, sein öffentliches Auftreten zu minimieren. Diesem Argument folgten oft Erklärungen, dass Proteste gefährlich seien und Navalny seine Anhänger in Gefahr bringe. Darüber hinaus wurde Navalny von den russischen Staatsmedien oft beschuldigt, Kinder und junge Erwachsene in seine politischen „Spiele“ einzubeziehen, zudem wurde er als „politischer Pädophiler“ bezeichnet (Uhov 2018).

(2) Navalny hat keine Agenda / keine Erfahrung als Politiker

„Aber wenn man darüber nachdenkt, tut er eigentlich nichts anderes, als Proteste und ‚Untersuchungen‘ zu organisieren, es wird nichts darüber gesagt, was er mit dem Staat tun wird.“ Diese Äußerung wurde von Putin ebenfalls getätigt. In einem Interview mit dem österreichischen Sender ORF warf Putin Navalny und der Opposition beispielsweise vor, keine positive Agenda und keine Vorschläge zur Lösung der Probleme – etwa der Korruptionen – zu haben (ORF 2018).

(3) Es wird sich nicht verändern oder sogar noch schlimmer werden, wenn Navalny an die Macht kommt

„Und Gott bewahre, dass er an die Macht kommt, denn Zeiten wie jetzt werden wie das Paradies erscheinen.“ Die politischen Institutionen sind in Russland generell schwach aufgestellt (Lassila 2016). Wie von Shlapentokh (2006), Stickley et al. (2009) und Tolochko (2019) unterstrichen wird, zählt das Vertrauen der Menschen in die politischen Institutionen in Russland zu den niedrigsten der Welt. Darüber hinaus weist Russland laut Transparency International (2019) eines der höchsten Korruptionsniveaus weltweit auf (Rang 138/180). Es ist daher nicht überraschend, dass dieses Argument ein gewisses Gewicht und eine starke Resonanz beim Publikum erzeugt.

(4) Die Ermittlungen liefern keine Beweise / Navalny ist ein Lügner

„Lesha, wie unterscheiden Sie sich von denen an der Spitze? Sie lügen auch, wo sind die wirklichen Beweise?“ Dieser Kommentar steht für die Zweifel oder die absichtliche Respektlosigkeit gegenüber den von Navalny durchgeführten Untersuchungen. Es wird behauptet, dass, wenn es möglich wäre, die Korruptionsvorhaben zu beweisen, Medwedew bereits zurückgetreten und von der Polizei verhört worden wäre. Es wurde auch oft erwähnt, dass Navalny vom Westen und der USA unterstützt werden würde – dies steht ebenfalls im Einklang mit den Propaganda-Botschaften.

Es gab zudem viele beleidigende Bezeichnungen, die weniger argumentativer Natur sind. Das Gleiche gilt auch für jene Kommentare, die beispielsweise den Protest oder die Protestierenden angreifen: „Ich schlage vor, dass jeder zum Protest am 26. geht und auf diejenigen pisst, die Navalny unterstützen.“ Folgende Themen wurden bei der Analyse der Kommentare, die die Proteste missbilligen, ebenfalls identifiziert:

(5) Die Proteste sind illegal und gefährlich

„Denken Sie an Ihre Kinder, bevor Sie [zu den Protesten] gehen. Nach der Verwaltungshaft werden sie nicht mehr arbeiten können.“ Nach dem Bolotnaja-Fall (2011–2012) wurde das russische Kundgebungsgesetz erheblich geändert, so wurden die Geldstrafen erhöht und die Möglichkeit von Festnahmen aufgrund der Teilnahme an „nicht genehmigten“ Protesten erleichtert. Die Verfassung der Russischen Föderation gewährt jedoch das Recht auf friedliche Kundgebungen und Versammlungen. In den Beiträgen wurde auf Grundlage dieses Widerspruchs argumentiert. Durch Aussagen wie „Ich habe Angst, gefeuert zu werden“ und „Es wird angenommen, dass alle Protestierenden erschossen werden“ wurden absichtlich Ängste geschürt.

(6) Die Proteste sind vergeblich

„Der Protest wird scheitern, mehrere hundert Menschen werden inhaftiert, niemand wird etwas erreichen.“ Wird die allgemeine Erosion der Protestaktivität in Russland nach der Verschärfung des Versammlungsgesetzes berücksichtigt, scheint dieses Argument die vorherrschende Stimmung zu reflektieren: Die Proteste in Russland finden keine Unterstützung in der Öffentlichkeit. Die Zivilgesellschaft steckt in einer tiefen Krise (Evans Jr. 2012). Die Protestierenden werden von den Behörden und den staatlichen Medien an den Rand gedrängt, zum Beispiel nannte der russische Staatsjournalist Vladimir Solov'ev diese in einem Interview für den staatlichen Radiosender „immerwährende 2 % der Scheiße“ (Business Online 2017).

(7) Wir wollen das ukrainische Szenario nicht wiederholen

„Sie werden Russland verschwenden und hier eine zweite Ukraine initiieren.“ Dies widerspricht dem Argument (6) und vergleicht die friedlichen Proteste mit Revolutionen, Staatsstreich, Massenunruhen und vor allem mit dem Euromaidan.<sup>9</sup> Laut dem Ukraine-Krisenmedienzentrum war zwischen 2014 und 2017 ein Drittel der Nachrichten im russischen Fernsehen der Ukraine gewidmet und davon waren mehr als 90 % negativ konnotiert (Vinogradov 2018). Auch Aufrufe zur Stabilität sowie Gründe und Beispiele, warum die Revolution um jeden Preis verhindert werden sollte, zählen zu dieser Themengruppe.

---

<sup>9</sup> Euromaidan bezeichnet eine Reihe von Protesten in der Ukraine in den Jahren 2013 und 2014.

## (8) Provokationen

„Bringt Waffen mit! Die Bullen werden auf euch schießen!“ In mehreren Kommentaren wurde unter anderem zur Radikalisierung der Proteste, zu Gewalt oder sogar zum Einsatz von Waffen aufgerufen.

Positive und unterstützende Antworten machen etwa ein Drittel der Antworten aus (siehe Abbildung 1) und beinhalten folgende Unterthemen: allgemeine Unterstützung von Navalny bzw. seiner Organisation, Unterstützung des Protestes, Aufrufe zur Verbreitung der Informationen über den Protest, organisatorische Fragen, Proteststrategien, Offenlegung von Trollen und Gedichte zur Unterstützung des Protestes. Zudem wurde die Kategorie „Top“ aufgenommen, die meist bedeutungslose Kommentare („Kommentar für Top!“, „bitte mindestens fünf Kommentare schreiben!“), die mit dem Ziel gepostet wurden, eine rasch ansteigende Aktivität zu erzeugen, beinhaltet. Diese Kommentare werden verfasst, damit das Video in die Kategorie „Trending“ der YouTube-Startseite kommt, wodurch eine größere Zahl an Nutzerinnen und Nutzer erreicht wird.

1 % oder weniger der Kommentare unterstützten Putin, Medwedew oder die Regierung (siehe Abbildung 1), wobei diese vor allem die Erfahrung Putins als Politiker und seinen Erfolg in der Außenpolitik unterstrichen: „[...] Russland braucht ihn als starken Politiker“. 10 %, 15 % und 5 % der Kommentare für die Videos 1 bis 3 waren negativ. Sie verurteilten in erster Linie die Korruption und beschuldigten Putin und Medwedew, diese nicht zu bekämpfen, sondern zu unterstützen: „Der Putler ist an allem schuld. Er ließ zu, dass jeder frei stehlen und Land und Leute ausrauben konnte.“

## 4.2 Strategien des ideologischen Trolling

In dieser Phase wurden jene Kommentare analysiert, die von den Konten generiert wurden, die von anderen Benutzerinnen und Benutzern als bezahlte Trolle (N = 127) identifiziert wurden. Abbildung 2 stellt die Verteilung der Strategien dar. Wenn die Benutzerin oder der Benutzer mehr als einen Kommentar abgegeben hatte, wurde die Strategie ermittelt, die am häufigsten verwendet wurde.

### (1) Förderung / Lob / Cheerleading

„Putin hat die Erhöhung des Rentenalters für Frauen reduziert! Jetzt 60 Jahre statt 63! Ich bin so stolz!“ In diesem Fall war es die am wenigsten angewandte Strategie (9,4 %). Die Regierungsagenda wurde hier argumentativ und nicht-argumentativ unterstützt. Zu den Themen

zählten die Preisung Putins und seiner Qualitäten („mächtiger Präsident“), die allgemeine Situation in Russland im Vergleich zu den 1990er-Jahren und eine Beschreibung der Verbesserungen der Wirtschaft und des Wohlstands („Jeder hat jetzt eine Wohnung und eine Datscha“).

(2) Unterbrechung / „Whataboutismus“

Um die Diskussion zu stören, verlagern die Trolle häufig die Diskussion auf die Situation in anderen Ländern. Sie verglichen etwa die Proteste, andere Ereignisse oder die allgemeine Situation mit Maidan, die in der Ukraine stattfand, mit Pariser Kundgebungen oder mit den USA: „Schauen Sie sich die Proteste in den USA an, was passiert, wenn der Polizist angegriffen wird? Sofort gefickt wird, wer immer das tut ... also viel Glück!“, „... wiederholen Sie nicht die Geschichte der Ukraine, denken Sie nach!“

(3) Respektlosigkeit / Aggression / Beleidigungen

Dieser Hauptmodus beinhaltet nicht-argumentative, aggressive, feindselige oder emotionale Respektlosigkeit gegenüber Navalny, seinen Anhängerinnen und Anhängern, Protesten und Demonstrantinnen und Demonstranten. Diese Strategie macht etwa ein Drittel der Äußerungen aus, etwa durch Übergriffe wie „[Navalny], du bist eine anale Hure, häng dich schon mit deinen dummen analen Anhängern auf!“ Dazu zählt auch, dass der Familienname Navalny in verschiedene anstößige Spitznamen umgewandelt wurde: Analny, Navalny – aus dem Wort Lüge, Nakralny – aus dem Wort stehlen, Carnivaly – aus dem Wort Karneval. Dies deckt sich mit Putins Weigerung, Navalny bei seinem Namen zu nennen, und ihn stattdessen als „diese Person“ oder „diesen Herrn“ zu bezeichnen (Robinson 2018).

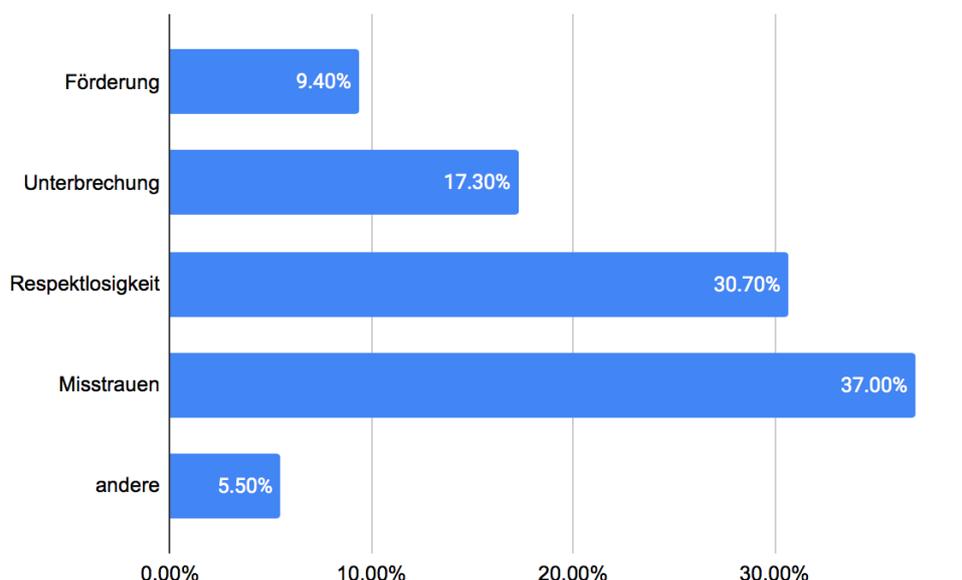


Abbildung 2: Vier Hauptmodi der Trolle (eigene Darstellung)

#### (4) Misstrauen

Misstrauen erwies sich als die beliebteste Strategie. Hierzu zählen Diskussionen mit dem Ziel, Misstrauen in Navalny als Politiker oder gegenüber den Protesten zu verbreiten. So finden sich hier einige ‚sachliche‘ Aussagen wie „die Unterstützung für Leschka beträgt landesweit nur 2 %“ oder welche zum Verbot der oppositionellen Kandidatinnen und Kandidaten für die Kandidatur zum Moskauer Stadtrat: „Sie hätten nicht mit der Datenbank betrügen sollen.“ Ein weiteres Argument war das allgemeine Misstrauen gegenüber Politikerinnen und Politikern und den Institutionen: „Alexej, Sie wollen Russland wirklich spalten, Sie sind ein Feind Russlands.“ Die Proteste wurden zudem als vergeblich, gefährlich für die Teilnehmerinnen und Teilnehmer oder ihre Angehörigen, als illegal, kindisch oder als politisches Spielzeug für Navalny dargestellt.

### 4.3 Identifizierung von Trollen

Nur 34 von 127 (27 %) Benutzerinnen und Benutzer verfügen über Daten, die den Zeitpunkt der Erstellung des Kontos angeben. Keines der Konten wurde unmittelbar vor den besprochenen Videos erstellt. Allerdings weisen 67 Benutzerinnen und Benutzer (54 %) keinerlei Aktivität auf ihren Kanälen auf. Jenen mit keiner oder geringer Aktivität wurde 1 Punkt zugewiesen. 29 Benutzerinnen und Benutzer (23 %) waren übermäßig aktiv (20 oder mehr Kommentare) beim Kommentieren der Videos und 31 (24 %) mäßig aktiv (5 oder mehr Kommentare). Das bedeutet, dass sich beinahe die Hälfte der Nutzerinnen und Nutzer aus der Stichprobe aktiv zu den Videos äußerte. Bei übermäßiger Aktivität erhielten sie einen ganzen Punkt und einen halben Punkt bei mäßiger Aktivität. 39 Benutzerinnen und Benutzer (31 %) gehörten zu den Erstkommentatorinnen und Erstkommentatoren und zehn weitere kommentierten das Video täglich in einem Zeitraum von mindestens vier Tagen (wurde noch 1 Punkt zugewiesen). Insgesamt konnten 17 Nutzerinnen und Nutzer mit hoher Sicherheit (2,5 bis 3 Punkte), 41 mit Sicherheit (1,5 bis 2 Punkte) und 41 weitere mit geringer Sicherheit (0,5 bis 1 Punkt) als bezahlte Trolle identifiziert werden.

## 5. Diskussion

Die von Alexej Navalny und seinem Team auf dessen YouTube-Plattform geschaffene Community ist aufgrund der hohen Resonanz beispiellos. Im Januar 2020 hatte sein Hauptkanal 3,2 Millionen Abonnentinnen und Abonnenten. Die dadurch geschaffene Community dient als unabhängiges Forum, in dem nicht nur seine Anhängerinnen und Anhänger, sondern auch die

Kritikerinnen und Kritiker einen offenen Dialog führen können. Die Plattform ist kein Gegenstand der Wahlgesetzgebung. Sie ist vom Staat unabhängig und existiert parallel zu den traditionellen russischen Medien, in denen die Meinungsvielfalt nicht gegeben ist – was in Russland weithin als wesentlicher Grund für eine funktionierende Demokratie akzeptiert wird. Der nichtkommerzielle Charakter der Plattform ist jedoch durchaus fragwürdig. Papacharissi (2008) plädiert für die Schaffung eines neuen hybriden Modells von Öffentlichkeiten, bei dem der kommerzielle Charakter der Plattform die bürgerliche Rhetorik noch zulässt. Die Kanäle von Navalny sind jedoch weder Bestandteil eines Partnerschaftsprogramms auf YouTube, noch in sonstige kommerzielle Prozesse eingebunden. Die Arbeit wird zu einem großen Teil durch freiwillige Spenden der Unterstützerinnen und Unterstützer finanziert. Informationen über die Proteste werden ausgetauscht, Verhaltensstrategien und notwendige Schritte im Fall von Polizeiverstößen diskutiert und darüber hinaus die Proteste live auf dem Kanal übertragen.

Die Plattform wurde neben den Facebook-Gruppen, die von den lokalen Protestorganisationen geführt werden, zum wichtigsten Protest-Organisationsnetzwerk in Russland. Dies sind die realen politischen Auswirkungen, die sich in der Schaffung der von Fraser (1990) kritisierten Gesetzgebung niederschlagen könnten. Indem die Reaktionen der Nutzerinnen und Nutzer auf die Videos zu drei verschiedenen Protesten analysiert wurden, konnte gezeigt werden, wie sie in einen Dialog über die Unterstützung oder Missbilligung von Navalny und die Ursache des Protestes getreten sind. Selbst die negativen Reaktionen waren zum größten Teil argumentativer Natur und diskutierten durchaus berechtigte Fragen hinsichtlich der Gefahr der Proteste, ihrer Bedeutung und ihres Zwecks, der politischen Agenda von Navalny, seiner Erfahrung, des Gerichtsverfahrens gegen ihn und seinen Bruder sowie der möglichen Zukunft im Fall einer Machtübernahme. Das Engagement und die Vielfalt der Meinungen (z.B. haben in Video 1 die Unterstützung und die Kritik gegenüber Navalny fast gleich große Anteile) entsprechen der Beschreibung der habermasianischen Öffentlichkeit.

Was jedoch Reflexivität, Rollenübernahme, respektvolles Zuhören und Aufrichtigkeit (vgl. Dahlberg 2001) auf YouTube betrifft, gibt es gewisse Bedenken. Milliken et al. (2008) stellen etwa fest, dass die Kommentare zu den Videos zumeist negativer Natur sind. Es kommt nicht selten vor, dass die anonymen Kommentare das Fluchen als „eine Zurschaustellung von (absichtlicher) Unhöflichkeit enthalten, die dem verbalen Missbrauch, der instrumentellen Aggression und dem Aufbau von Machthierarchien unter anonymen Internetsurfern dient“ (Dyner 2012: 27). Auch in der vorliegenden Untersuchung wurden zahlreiche Beleidigungen, der Gebrauch vulgärer Sprache, Hassreden und Aufrufe zu Gewalt gegenüber Navalny, Putin und ihren jeweiligen Anhängerinnen und Anhängern identifiziert.

Neben der Feindseligkeit wird das ideologische *Astroturfing* aufgrund seiner manipulativen Macht zu einem ernstem Problem. Die zentrale Aufgabe der Forscherinnen und Forschern besteht darin, die Ziele der an der Diskussion teilnehmenden Personen zu entschlüsseln – geben diese ihre wahren Überzeugungen wider oder wurden sie dafür bezahlt, den Anschein einer vorherrschenden Meinung zu erwecken? Da die Daten nicht zeitgleich gesammelt wurden, gab es keine Möglichkeit, zu ermitteln, wie viele Kommentare und Konten in der Zwischenzeit gelöscht wurden. Dadurch ist die Stichprobe lediglich auf die gespeicherten Kommentare beschränkt. Aus dem vorliegenden Beispiel wurde eine Stichprobe an Benutzerinnen und Benutzern extrahiert, die von der Community als bezahlte Trolle bezeichnet wurden. Es hat sich gezeigt, dass sich die Trollstrategie in Russland von jener in China, wo Cheerleading und das argumentative und nicht-argumentative Lob die Hauptstrategien sind, unterscheidet (vgl. King et al. 2017). Denn es konnten kaum eine Förderung und Unterstützung der Regierungsgagenda oder -maßnahmen festgestellt werden. Der Prozentsatz war allerdings höher als in der gesamten allgemeinen Stichprobe.

King et al. (2017) zufolge besteht das Hauptprinzip der *Astroturfer* darin, sich nicht an kontroversen Diskussionen zu beteiligen. Dieses scheint bei den analysierten Videos jedoch nicht der Fall zu sein, da selbst die Botschaften, die Putin fördern und unterstützen, einen eher kritischen Ton haben. Zum Beispiel: „Das russische Volk braucht eine starke Führung“, „Es ist nicht Putin, sondern sein liberales Team, das die Rentenreform vorgeschlagen hat“. Das Argument „Stabilität“ war in den Kommentaren zur Unterstützung Putins und des Regimes erkennbar, aber auch in denen, die Navalny und die Proteste anprangerten. Nach Tolochko (2019) ist das Vertrauen in den Präsidenten traditionell viel höher als in andere politische Institutionen. Das kann als Erklärung dafür dienen, warum Putin häufig freigesprochen und sein Team für die unpopulären Reformen verantwortlich gemacht wird. Es wurde auch davon ausgegangen, dass Aufrufe zum Handeln oder zur Provokation anderer Nutzerinnen und Nutzer, wie z.B. die Aufforderung, Waffen zu den Protesten zu bringen, den größten Teil der Kommentare ausmachen werden. Allerdings entsprechen nur 1 % mäßig provozierender Botschaften dieser Strategie und so wurde diese nicht in das endgültige Codebuch aufgenommen.

Die Strategien der bezahlten Trolle konzentrieren sich darauf, Respektlosigkeit zu zeigen und Misstrauen zu schüren. Die Verunglimpfung Navalnys entspricht der von Putin selbst angewandten Vorgehensweise – er spricht in den offiziellen Interviews oder Erklärungen des Pressesekretärs Dmitri Peskow niemals den Familiennamen Navalnys aus und wertet die Rolle der „systemfremden“ oder liberalen Opposition, die Proteste sowie die Protestierenden ab. Die Proteste werden von den staatlichen russischen Fernsehsendern oft nicht erwähnt. Die offizielle Zahl der Protestierenden unterscheidet sich zumeist von der Zahl der von unabhängigen

Organisationen gemeldet (z. B. „Weißer Zähler“). Putin selbst vermeidet es, über die Proteste zu sprechen, oder vergleicht sie mit den Massenunruhen in der Ukraine. Die staatlichen Medien zweifeln an den Fähigkeiten der „Navalny-fanterie“, ein „starker“ politischer Anführer zu werden, wobei sie ihn oft als Kriminellen bezeichnen. Die *Astroturfer* greifen diese Argumente auf, um das Misstrauen in seine Handlungen, seine Agenda und seine Untersuchungen zu schüren.

Eine explizite Zensur scheint die Menschen zu verärgern und das Vertrauen in die Medien und die politischen Institutionen zu reduzieren. Darüber hinaus „[...] erfordert internationale Legitimität öffentliche Demonstrationen der Opposition“ (Robertson 2009: 546). Daher sind *Astroturfer* zu einem mächtigen Werkzeug geworden. Das Aufspüren von Trollen erfordert neue Strategien und umfangreiche Untersuchungen. Aus der Matrix, die in dieser Forschung erstellt wurde, konnten nur 13 % der Konten mit hoher Wahrscheinlichkeit als bezahlte Trolle klassifiziert werden. Die von den Kommentatorinnen und Kommentatoren und bezahlten Trollen verwendeten Argumente konnten mit den von der offiziellen Propagandamaschinerie verbreiteten Nachrichten abgeglichen werden. Darüber hinaus wiederholen einige Argumente und Kommentare wortgetreu Putins Ankündigungen aus seinen offiziellen Interviews. Dieser Befund kann aber nicht als Unterscheidungsmerkmal zwischen den regelmäßigen Nutzerinnen und Nutzern und den Trollen verwendet werden, da die regelmäßigen Nutzerinnen und Nutzer diese Nachrichten ebenso aufgreifen und sie ohne böswillige Absicht verbreiten könnten. Dies erklärten jedoch auch die Strategien, die von den Trollen angewandt werden. Nicht nur die allgemeine Öffentlichkeit ist mit diesen Botschaften vertraut, auch der Regierung wird die Verbreitung dieser dadurch erleichtert – etwa indem sie einen Pool von Äußerungen erstellt, die die offizielle Agenda unterstützen.

## 6. Fazit

Diese Forschungsarbeit befasst sich mit der konzeptionellen Frage der virtuellen Öffentlichkeit und der methodischen Frage der Identifizierung von Trollen. Im Kontext Russlands, wo das Vertrauen in die staatlichen Institutionen und die öffentlichen Medien im Allgemeinen gering ist und weiter abnimmt, ist es sehr wichtig, die virtuelle Öffentlichkeit in den Diskurs einzubinden. Trotz der Fragmentierung des Online-Publikums sind die neuen sozialen Medien die Plattformen, auf denen demokratische Prozesse erkannt werden können. Einer der Kritikpunkte hinsichtlich der Ideen Habermas' wurde von Fraser (1990) geäußert: die Möglichkeit der Öffentlichkeit, die Mittel für demokratisches Handeln bereitzustellen. Im Fall des YouTube-Kanals von Navalny zeigt sich, dass diese Mittel zur Verfügung gestellt werden.

Nach der Übernahme von *vk.com* im Jahr 2014 und dem Verbot von Telegram im Jahr 2018 sind die wichtigsten oppositionellen Kräfte Russlands auf YouTube umgezogen. Da YouTube nicht nur Unterhaltungsinhalte anbietet, hat die Plattform schnell an gesellschaftspolitischer Bedeutung gewonnen (vgl. Stross 2008). Sie bietet politischen Dissidentinnen und Dissidenten die Möglichkeit, unpopuläre Meinungen zu äußern, die ohne staatliche Medien verbreitet werden können. Auf YouTube können zudem Netzwerke von Unterstützerinnen und Unterstützern aufgebaut werden, die durchaus Protestpotenzial haben. Die Erforschung von YouTube in Russland ist von entscheidender Bedeutung, da diese Plattform – im Gegensatz zu sozialen Medien wie Twitter und Facebook – in diesem Land einen hohen Marktanteil hat.

Diese Untersuchung liefert eine Basis für die Etablierung einer Matrix von Kriterien, die auf der Grundlage anderer Attribute erstellt wurde, um mit einer gewissen Wahrscheinlichkeit identifizieren zu können, ob die Benutzerin oder der Benutzer tatsächlich ein *Astroturfer* ist. Diese Forschung konzentriert sich auf die manipulativen Techniken, die die Existenz einer virtuellen Öffentlichkeit in Frage stellen. *Astroturfer* bzw. bezahlte Trolle sind ein Phänomen, das auch in der Wissenschaft immer mehr thematisiert und untersucht wird. So wurde – mit unterschiedlichen Ergebnissen – in mehreren Studien versucht, bezahlte Trolle zu identifizieren.

Eine Hürde in der Forschung stellt das Fehlen von *Ground-Truth*-Daten dar, die Effizienz dieser Kriterien kann somit nicht überprüft werden. Dennoch zeigt die Forschung, dass Trolle ein organisiertes, ähnliches Verhalten aufweisen, was auf eine Koordination bzw. das Kopieren von Verhalten schließen lässt. Das Belegen von ähnlichen Strategien, Aussagen, Wörtern und Argumentationen ist ein wichtiger erster Schritt zur Identifizierung von Trollen. Dieser Artikel kann somit einen Grundstein für weitere Untersuchungen der öffentlichen Sphäre auf YouTube und der Störung der virtuellen Öffentlichkeit durch die Beteiligung bezahlter Trolle bilden. Zukünftige Forschungsarbeiten könnten die Netzwerkanalyse als alternative Methode zur Extraktion von Trollproben nutzen und die Aktivitäten einer tieferen Analyse unterziehen.

## Literatur

Ananyev, Maxim/Sobolev, Anton (2017): *Fantastic Beasts and Whether They Matter: Do Internet ‘Trolls’ Influence Political Conversations in Russia*. Paper presented at the Meeting of the Midwest Political Science Association, April 2017, Chicago, IL.

Antony, Mary Grace/Thomas, Ryan (2010): „This is citizen journalism at its finest“: YouTube and the public sphere in the Oscar Grant shooting incident. In: *New Media & Society* 12(8), 1280-1296.

Aro, Jessikka (2016): *The cyberspace war: propaganda and trolling as warfare tools*. In: *European view* 15(1), 121-132.

- Badawy, Adam/Emilio Ferrara/Kristina Lerman (2018): Analyzing the digital traces of political manipulation: The 2016 russian interference twitter campaign. In: IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 258-265.
- Bodnar, Adam/Hamilton, Michael/Jarman, Neil/Martynenko, Oleg/Ostaf, Serghei/Zhovtis, Yevgeniy (2013): May 6 2012 Events on Bolotnaya Square in Moscow: Expert Evaluation. Online: [https://ueaeprints.uea.ac.uk/id/eprint/48443/1/Hamilton\\_May\\_6\\_2012.pdf](https://ueaeprints.uea.ac.uk/id/eprint/48443/1/Hamilton_May_6_2012.pdf) (11.07.2020).
- Bolognani, Marta (2011): Rang de Basanti in Pakistan? Elite student Activism, the Emergence of a Virtual Globalized Public Sphere, and the 2007 Emergency. In: Bolognani, Marta/Lyon, Stephen (Hg.): Pakistan and its diaspora. New York: Palgrave Macmillan, 19-42.
- Bradshaw, Samantha/Howard Philip (2017): Troops, trolls and troublemakers: A global inventory of organized social media manipulation. Oxford: Oxford University Press.
- Business Online (2017): Vladimir Solovyov on rallies against corruption: "Children of corrupt officials went to the streets!" Online: <https://www.business-gazeta.ru/news/348637> (11.07.2020).
- Colleoni, Elanor/Rozza, Alessandro/Arvidsson, Adam (2014): Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. In: Journal of communication 64(2), 317-332.
- Dahlberg, Lincoln (2001): Computer-mediated communication and the public sphere: A critical analysis. In: Journal of Computer-mediated communication 7(1), 1-40.
- Dahlberg, Lincoln (2007): The Internet, deliberative democracy, and power: Radicalizing the public sphere. In: International Journal of Media & Cultural Politics 3(1), 47-64.
- Holmes, L. David (1997): Virtual politics: Identity and community in cyberspace. New York: Sage Publishing.
- Dynel, Marta (2012): Swearing methodologically: the (im)politeness of expletives in anonymous commentaries on Youtube. In: Journal of English studies 10, 25-50.
- Edgerly, Stephanie/Vraga, Emily/Fung, Timothy/ Moon, Tae Joon/Yoo Woo Hyun/Veenstra, Aaron (2009): YouTube as a public sphere: The proposition 8 debate. In: Association of Internet Researchers Conference, Milwaukee, Wi.
- Enikolopov, Ruben/Makarin, Alexey/Petrova, Maria (2020): Social Media and Protest Participation: Evidence from Russia. In: Econometrica 88(4), 1479-1514. DOI: <http://dx.doi.org/10.2139/ssrn.2696236>.
- European Parliament (2016): European Parliament resolution of 23 November 2016 on EU strategic communication to counteract propaganda against it by third parties (2016/2030(INI)). Online: [https://www.europarl.europa.eu/doceo/document/TA-8-2016-0441\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-8-2016-0441_EN.html) (11.07.2020).
- Evans Jr, Alfred (2012): Protests and civil society in Russia: The struggle for the Khimki Forest. In: Communist and Post-Communist Studies 45(3-4), 233-242.
- Fraser, Nancy (1990): Rethinking the Public Sphere: A contribution to the Critique of Actually Existing Democracy. In: Social Text, Nr. 25/26, 56-80. DOI: <https://doi.org/10.2307/466240>.

- Freedom House (2013): Freedom on the Net: A Global Assessment of Internet and Digital Media. Online: <https://freedomhouse.org/sites/default/files/resources/FOTN%202013%20Summary%20of%20Findings.pdf> (11.07.2020).
- Garnham, Nicholas (1990): Capitalism and communication: Global culture and the economics of in-formation. New York: Sage Publishing.
- Garnham, Nicholas (2007): Habermas and the public sphere. In: *Global Media and Communication* 3(2), 201-214.
- Geiger, R. Stuart (2009): Does Habermas understand the internet? The algorithmic construction of the blogo/public sphere. In: *A Journal of Communication, Culture, and Technology* 10(1), 1-29.
- Georgakopoulou, Alexandra/Spilioti, Tereza (2015): *The Routledge handbook of language and digital communication*. Abingdon: Routledge.
- Glazunova, Sofya (2020): Mediating Populist Discourse in Russia via YouTube: The Case of Alexey Navalny. In: Akimov, Alexandr/Kazakevitch, Gennadi (Hg.): *30 Years since the Fall of the Berlin Wall*. Singapore: Palgrave Macmillan, 191-213.
- Godin, Paul/Davies, Jacqueline/Heyman, Bob/Reynolds, Lisa/Simpson, Alan/Floyd, Mike (2007): Open-ing communicative space: A Habermasian understanding of a user-led participatory research project. In: *The Journal of Forensic Psychiatry & Psychology* 18(4), 452-469.
- Habermas, Jürgen (1981): *Theorie des kommunikativen Handelns*. Frankfurt: Suhrkamp, 1049-1054.
- Habermas, Jürgen (1989): The Public Sphere: An Encyclopedia Article. In: Brinner, Stephen E./Kellner, Douglas (Hg.): *Critical Theory and Society. A Reader*. New York: Routledge, 136-142.
- Habermas, Jürgen (2006): Political communication in media society: Does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research. In: *Communication theory* 16(4), 411-426.
- Han, Rongbin (2015): Manufacturing Consent in Cyberspace: China's "Fifty-Cent Army". In: *Journal of Current Chinese Affairs* 44(2), 105-134.
- Hardaker, Claire (2013): "Uh.... not to be nitpicky, but... the past tense of drag is dragged, not drug.": An overview of trolling strategies. In: *Journal of Language Aggression and Conflict* 1(1), 58-86.
- Hussain, Muhammad Nihal/Tokdemir, Serpil/Agarwal, Nitin/Al-Khateeb, Samer (2018): Analyzing disinformation and crowd manipulation tactics on YouTube. In: *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 1092-1095.
- Jacobs, N. Ronald (1999): Race, media and civil society. In: *International Sociology* 14(3)s, 355-372.
- Johnson, Pauline (2006): *Habermas: Rescuing the public sphere*. New York: Routledge.
- Kellner, Douglas (2000): Habermas, the public sphere, and democracy: A critical intervention. In: *Perspectives on Habermas* 1(1), 259-288.

- King, Gary/Pan, Jennifer/Roberts, E. Margaret (2017): How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. In: *American Political Science Review* 111(3), 484-501.
- Kirby, Jen (2018): What to know about the Russian troll factory listed in Mueller's indictment. In: [vox.com](https://www.vox.com), 16.02.2018. Online: <https://www.vox.com/2018/2/16/17020974/mueller-indictment-internet-research-agency> (11.07.2020).
- Lassila, Jussi (2016): Aleksei Naval'nyi and populist re-ordering of Putin's stability. In: *Europe-Asia Studies* 68(1), 118-137.
- Linville, Darren L./Warren, L. Patrick (2019): Touched by the Trolls: How and Why a Coordinated In-formation Operation Interacts with Outsiders. Online: <https://pdfs.semanticscholar.org/deba/83d06283496135f5284e86201039ead352d3.pdf> (11.07.2020).
- Logan, Kelty (2014): Why isn't everyone doing it? A comparison of antecedents to following brands on Twitter and Facebook. In: *Journal of Interactive Advertising* 14(2), 60-72.
- Miller, Blake (2016): Automated detection of Chinese government astroturfers using network and social metadata. Online: <http://dx.doi.org/10.2139/ssrn.2738325> (11.07.2020).
- Milliken, Mary/Gibson, Kerri/O'Donnell, Susan/Singer, Janice (2008): User-generated online video and the Atlantic Canadian public sphere: A YouTube study. In: *Proceedings of the International Communication Association Annual Conference (ICA 2008)*, Montreal, Quebec, Canada. Online: [https://www.researchgate.net/profile/S\\_Odonnell/publication/44068095\\_User-generated\\_Online\\_Video\\_and\\_the\\_Atlantic\\_Canadian\\_Public\\_Sphere\\_A\\_YouTube\\_Study/links/02e7e5379f25f3d01a000000/User-generated-Online-Video-and-the-Atlantic-Canadian-Public-Sphere-A-YouTube-Study.pdf](https://www.researchgate.net/profile/S_Odonnell/publication/44068095_User-generated_Online_Video_and_the_Atlantic_Canadian_Public_Sphere_A_YouTube_Study/links/02e7e5379f25f3d01a000000/User-generated-Online-Video-and-the-Atlantic-Canadian-Public-Sphere-A-YouTube-Study.pdf) (11.07.2020).
- Mihaylov, Todor/Georgiev, Georgi/Nakov, Preslav (2015): Finding opinion manipulation trolls in news community forums. In: *Proceedings of the nineteenth conference on computational natural language learning*, 310-314.
- NTIA (1995): *Falling through the net: A survey of the "Have Nots" in rural and urban America*. Washington: National Telecommunications and Information Administration, U.S. Department of Commerce. Online: <http://www.ntia.doc.gov/ntiahome/fallingthru.html> (11.07.2020).
- ORF (2018): Keinerlei Zweifel an der Demokratie. 04.06.2018. Online: <https://orf.at/v2/stories/2441362/2441364/> (11.07.2020).
- Papacharissi, Zizi (2002): The virtual sphere: The internet as a public sphere. In: *New Media & Society* 4(1), 9-27.
- Papacharissi, Zizi (2008): The virtual sphere 2.0: The Internet, the public sphere, and beyond. In: *Routledge Handbook of Internet Politics*. London: Routledge, 246-261.
- Peng, Jian/Choo, Raymond Kim-Kwang/Ashman, Helen (2016): Astroturfing detection in social media: Using binary n-gram analysis for authorship attribution. 2016 IEEE TrustCom/BigDataSE/ISPA, Tianjin, 121-128. DOI: <https://www.doi.org/10.1109/TrustCom.2016.0054>.
- Rieder, Bernhard (2015): YouTube Data Tools (Version 1.11) [Software]. Online: <https://tools.digitalmethods.net/netvizz/youtube/> (14.08.2020).

- Robertson, Graeme B. (2009): *Managing Society: Protest, Civil Society, and Regime in Putin's Russia*. In: *Slavic Review* 68(3), 528-547.
- Robinson, Adam (2018): Russian opposition leader Navalny's name shifts snow in Moscow. In: *bbc.com, News*, 12.02.2018. Online: <https://www.bbc.com/news/world-europe-43034642> (14.08.2020).
- Sanovich, Sergey (2017): *Computational Propaganda in Russia: the Origins of Digital Misinformation*. Working Paper No. 2017.3. In: Woolley, Samuel/Howard, N. Philip (Hg.): *Computational Propaganda Worldwide*. Project on Computational Propaganda. Oxford: University of Oxford Press. Online: <https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Comprop-Russia.pdf> (14.08.2020).
- Seah, Chun Wei/Chieu, Hai Leong /Chai, Kian Ming/Loo-Nin Teow/Yeong, Lee Wei (2015): *Troll detection by domain-adapting sentiment analysis*. 18th International Conference on Information Fusion (Fusion), Washington, DC, 792-799.
- Sobolev, Anton (2019): *How Pro-Government „trolls“ Influence Online Conversations in Russia*. Online: <http://www.wpsanet.org/papers/docs/2019W-Feb-Anton-Sobolev-Trolls-VA.pdf> (17.08.2020).
- Shlapentokh, Vladimir (2006): *Trust in public institutions in Russia: The lowest in the world*. In: *Communist and Post-Communist Studies* 39(2), 153-174.
- Sparks, Colin (2001): *The Internet and the global public sphere*. In: Bennett, Lance/Entman, Robert (Hg.): *Mediated Politics: Communication in the Future of Democracy*, Cambridge: Cambridge University Press, 75-95.
- Stickley, Andrew/Ferlander, Sara/Jukkala, Tanya/Carlson, Per/Kislitsyna, Olga/Mäkinen, Ilkka Henrik (2009): *Institutional Trust in Contemporary Moscow*. In: *Europe-Asia Studies* 61(5), 779-796.
- Strauss, Anselm/Corbin, Juliet (1994): *Grounded Theory Methodology*. In: *Handbook of Qualitative Research* 17(1), 273-285.
- Stross, Randall (2008): *Planet Google: One Company's Audacious Plan to Organize Everything We Know*. New York: Simon and Schuster.
- Stukal, Denis/Sanovich, Sergey/Bonneau, Richard/Tucker, A. Joshua (2017): *Detecting Bots on Russian Political Twitter*. In: *Big Data* 5(4), 310-324.
- The Moscow Times (2014): *Vkontakte Founder Pavel Durov Learns He's Been Fired Through Media*. 22.04.2014. Online: <https://www.themoscowtimes.com/archive/vkontakte-founder-pavel-durov-learns-hes-been-fired-through-media> (17.08.2020).
- Thomas, Graham/ Wyatt, Sally (1999): *Shaping cyberspace – Interpreting and transforming the Internet*. In: *Research Policy* 28(7), 681-698.
- Tomaiuolo, Michele/Lombardo, Gianfranco/Mordonini, Monica/Cagnoni, Stefano/ Poggi, Agostino (2020): *A Survey on Troll Detection*. In: *Future Internet* 12(31).
- Tolochko, Aleksandr (2019): *Trust in political institutions as a resource of reputation capital of the state*. In: *Revista San Gregorio* 36, 6-18.
- Uhov, Ilia (2018): *Political paedophiles, or why Navalny will again escape responsibility*. In: *Life*, 10.05.2018. Online: <https://life.ru/p/1115407> (17.08.2020).

- Valtysson, Bjarki (2012): Facebook as a Digital Public Sphere: Processes of Colonization and Emancipation. In: TripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society 10(1), 77-91.
- Vinogradov, Oleksij (2018): „Un-country“ and „civil war“: What is Russian propaganda about Ukraine based on? In: Radio Svoboda, 12.09.2018. Online: <https://www.radiosvoboda.org/a/29595896.html> (17.08.2020).
- Volchek, Dmitrij (2018): Die Verlierer arbeiten als Kremlinbots: Die traurigen Wochentage der Trollfabrik. In: Svoboda, 06.03.2020. Online: <https://www.svoboda.org/a/29079871.html> (17.08.2020).
- Walzer, Michael (1999): On Toleration. Yale: Yale University Press.
- Zhang, Jerry/Carpenter, Darrell/Ko, Myung (2013): Online astroturfing: A theoretical perspective. Online: [https://www.researchgate.net/publication/286729041\\_Online\\_astroturfing\\_A\\_theoretical\\_perspective/link/56df1c4208ae979addef5188/download](https://www.researchgate.net/publication/286729041_Online_astroturfing_A_theoretical_perspective/link/56df1c4208ae979addef5188/download) (17.08.2020).
- Zhdanova, Mariia/Orlova, Dariya (2017): Computational Propaganda in Ukraine: Caught between external threats and internal challenges. Working Paper No. 2017.9. University of Oxford Press. Online: <https://blogs.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Comprop-Ukraine.pdf> (17.08.2020).
- Zelenkauskaitė, Asta/Niezgoda, Brandon (2017): Stop Kremlin trolls: Ideological trolling as calling out, rebuttal, and reactions on online news portal commenting. In: First Monday 22(5). Online: <https://firstmonday.org/article/view/7795/6225> (17.08.2020).

## Kurzbiografie der Autorin



**Yulia Belinskaya, MA**, Forschungsassistentin an der Universität Wien. Sie hat ihren ersten Master-Abschluss an der Technischen Universität St. Petersburg und ihren zweiten an der Universität Stockholm mit dem Schwerpunkt Medien und Kommunikation erhalten. In ihrem Dissertationsprojekt untersucht sie Legitimationsmechanismen restriktiver Mediengesetze. Ihre Forschungsinteressen umfassen Fragen der Meinungsfreiheit, städtischer Kommunikationsräume, der öffentlichen Sphäre und der digitalen Medien.

Kontakt: [yulia.belinskaya@univie.ac.at](mailto:yulia.belinskaya@univie.ac.at)