

Sourcing and Automation of Political News and Information During Three European Elections

Lisa-Maria Neudert¹, Philip Howard¹ ,
and Bence Kollanyi²

Social Media + Society
July-September 2019: 1–13
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2056305119863147
journals.sagepub.com/home/sms


Abstract

Voters increasingly rely on social media for news and information about politics. But increasingly, social media has emerged as a fertile soil for deliberately produced misinformation campaigns, conspiracy, and extremist alternative media. How does the sourcing of political news and information define contemporary political communication in different countries in Europe? To understand what users are sharing in their political communication, we analyzed large volumes of political conversation over a major social media platform—in real-time and native languages during campaign periods—for three major European elections. Rather than chasing a definition of what has come to be known as “fake news,” we produce a grounded typology of what users actually shared and apply rigorous coding and content analysis to define the types of sources, compare them in context with known forms of political news and information, and contrast their circulation patterns in France, the United Kingdom, and Germany. Based on this analysis, we offer a definition of “junk news” that refers to deliberately produced misleading, deceptive, and incorrect propaganda purporting to be real news. In the first multilingual, cross-national comparison of junk news sourcing and consumption over social media, we analyze over 4 million tweets from three elections and find that (1) users across Europe shared substantial amounts of junk news in varying qualities and quantities, (2) amplifier accounts drive low to medium levels of traffic and news sharing, and (3) Europeans still share large amounts of professionally produced information from media outlets, but other traditional sources of political information including political parties and government agencies are in decline.

Keywords

political news and information, elections, junk news, Europe, grounded typologies, news sources

Introduction

Voters have access to an immense amount of information during elections, but it is not always clear that they are accessing quality information, in efficient ways, when they come to decide how to vote (Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2018; Woolley & Howard, 2017). Recently, the viral spread of misinformation has emerged as a central concern on the public agenda. Around the globe, a wide range of politically and economically motivated actors take advantage of social media platforms to inject public discourse with computational propaganda, conspiracy, and extremist rhetoric. In an effort to undermine the integrity of political processes hostile foreign actors, radical groups at the fringe of the political spectrum and profit-driven alternative outlets rely on a manipulative amalgam of deception, falsehoods, and misleading sensationalism designed to manipulate. Amid

calls for regulation of social media platforms tasked with countering the spread of hyper-polarizing and hateful rhetoric, there is a heightened concern about the effect of misinformation on democracy (Bradshaw, Neudert, & Howard, 2018; Persily, 2017; Tucker, Theoharis, Roberts, & Barberá, 2017; Wardle & Derakhshan, 2017).

Efforts to evaluate the quality of the political news and information shared during elections—especially over social media—face a number of methodological and epistemic

¹University of Oxford, UK

²Corvinus University of Budapest, Hungary

Corresponding Author:

Philip Howard, Oxford Internet Institute, University of Oxford,
1 St Giles', Oxford OX1 3JS, UK.
Email: philip.howard@oii.ox.ac.uk



challenges. A growing number of scholars have presented empirical evidence about how large volumes of incorrect information spread over social media. However, scholarly efforts to conceptualize phenomena in relation to the viral spread and weaponization of disinformation have remained contentious, lacking scholarly consensus on definition and nomenclature. Specifically, “fake news” has proved difficult to operationalize, and despite in common parlance now, it is more of a political accusation than concept for evaluating news quality.

Understanding what political news and information voters are sharing must be the starting point for any academic discussion—and public conversation—about the impact of social media on public life. To address this shortcoming and to understand the salient qualities of political news and information on social media, this article puts forward a grounded typology of sources shared on social media. We collected 4.3 million tweets from three European elections in France, Germany, and the United Kingdom. Users in Europe posted links to 5,158 distinct domains which informed our systemic analysis. Drawing from political communication theory, we developed a grounded typology of political news and information. We find that users across Europe shared substantial amounts of junk news in varying qualities and quantities, but that large amounts of professionally produced news and information from media outlets dominated the traffic on social media by far. However, other traditional sources of political news and information are in decline. Relatively small amounts of news and information were generated by political parties and governments. Our analysis shows that traffic around the elections was predominantly organic and amplifier accounts drove only low to medium levels of shares.

Our grounded typology categorized sources of political news and information on a domain level, rather than assessing the veracity of individual stories. This approach accommodated for a comprehensive analysis of the aggregate reporting of web sources as a whole. Over our analysis of three elections, we consistently found a salient category of “junk news” purporting to be real political news and information in varying amounts. Our methodology describes “junk news” as sources that meet three out of the following five criteria: (1) *professionalism*, where sources fall short of the standards of professional journalistic practice including information about authors, editors, and owners; (2) *style*, where inflammatory language, suggestion, ad hominem attacks and misleading visuals are used; (3) *credibility*, where sources report on unsubstantiated claims, rely on conspiratorial and dubious sources, and do not post corrections; (4) *bias*, where reporting is highly biased and ideologically skewed and opinion is presented as fact; and (5) *counterfeit*, where sources mimic established news publications in branding, design, and content presentation.

For scholarship on political communication to advance, it is crucial to form both an empirical and theoretical understanding of political news and information on social media

during elections and how this content is shared. For each country in this comparative study, we have three research questions.

- Which political parties dominated Twitter conversation during their national elections?
- What are the salient quantities and qualities of sources of political news and information users shared over social media, and how much of this content was junk news?
- To what extent was political conversation on social media driven by inauthentic amplifier accounts as opposed to organically shared human traffic?

First, we develop a distinct conceptual framework of “junk news.” Second, we situate the foundational role of typology building for understanding new phenomena in political communication literature. Third, we discuss our methodology for developing a grounded typology of political news and information that were shared over Twitter during elections in France, the United Kingdom, and Germany. Fourth, we analyze patterns and trends across the public discourse across these three elections online. To conclude, we discuss the implications of our findings for both democracy and political communication research.

Junk News in Contemporary Political Discourse

Disinformation and specifically the spread of “fake news” have emerged as a prominent public issue in Europe (Marchal, Kollanyi, Neudert, & Howard, 2019). In the aftermath of the Brexit referendum and the 2016 Presidential elections in the United States, a growing number of scholars have become concerned with “fake news” phenomena. Political communication scholarship has become devoted to the spread of fake news over online networks and messenger applications, its effects on human behavior and cognition, its relationship with digital content production, and its impact on public discourse and democracy (Del Vicario et al., 2016; Machado, Kira, Narayanan, Kollanyi, & Howard, 2019; Vicario et al., 2016; Vosoughi, Roy, & Aral, 2018). Despite a growing body of literature on the phenomenon, a consistent definition of fake news has yet to emerge. The current operationalization of the term in political communication scholarship overwhelmingly lacks conceptual clarity and a comprehensive, grounded framework (Neudert & Marchal, 2019). A wide range of information maladies phenomena including malicious information operations, misleading headlines, alternative media outlets, and “click-bait” are referred to as fake news in popular parlance.

In light of the diverse scope of the fake news term, scholars have identified a range of epistemological challenges surrounding its definition, including its uncritical adoption in popular verbatim, and its appropriation through political actors (European Commission, 2018; Neudert, 2017). To

address these shortcomings, we have developed a conceptual framework that is grounded in real world data about what users are actually sharing on social media (on junk news conceptualization). As opposed to other attempts of operationalizing “fake news,” our analysis is rooted in a rigorous social media data analysis of sharing behavior rather than phenomenological evidence and hypothesizing. We analyzed 5,158 sources to develop a comprehensive definition of the new phenomenon. Junk news fulfills at least three out of the five following criteria: (1) professionalism, (2) style, (3) credibility, (4) bias, and (5) counterfeit. Drawing on political communication literature, we identify an increasing prevalence of “junk news” that is distinct from the range of different forms of problematic information online.

Our methodology for classifying sources distinguishes the content of the domain as whole (i.e., style, bias, credibility), information about authors and the organization (i.e., professionalism), and the layout and design of the domain itself (i.e., counterfeit). A source needs to exhibit at least three of our criteria to be labeled as junk news. For example, partisan news organizations that report with an ideological angle (bias) would not be considered junk news unless they also failed on at least two of the other criteria.

Professionalism

Junk news domains that fail on the professionalism criteria purposefully refrain from providing clear information about real authors, editors, publishers, and owners, and do not publish corrections on debunked information.

Counterfeit

Some sources mimic established news reporting including fonts, branding, and design. Junk news is stylistically disguised as professional news with references to news wire services and credible sources, as well as headlines written in a news tone with date, time, and location stamps. In pronounced cases, outlets will copy logos and counterfeit domains.

Style

Style is concerned with the literary devices and language used throughout news reporting. Junk news employs propaganda techniques to systematically manipulate users for political purposes that appeal to emotions, rather than cognition, including emotional language with emotive expressions and symbolism, ad hominem attacks, misleading headlines, exaggeration, inflammatory language, unsafe generalizations, logical fallacies, moving images, and lots of pictures or mobilizing memes.

Bias

These junk news sources are highly biased and ideologically skewed in the presentation and selection of their reporting

and publish opinion pieces as news. They exhibit systematic differences in the mapping from facts to news reports conveying strikingly different accounts of what actually transpired to misrepresentation and selective reporting.

Credibility

Websites that fall short of credibility typically report on unsubstantiated claims and quote conspiratorial and dubious sources to verify their reporting. They do not vet their sources, fact-check, or consult multiple sources.

Typologies and New Modes of Political Communication

Content typologies are a foundational element in political research, and especially relevant for examining and explicating new phenomena, unanticipated problems, or rapid changes in social systems (Aronovitch, 2012; Howard & Hussain, 2013; Swedberg, 2018). Broad categories of political news and communication have remained widely persistent over the years, and have been transported across national, cultural, and technological contexts though a number of scholars have debated the applicability of traditional frameworks to new media (see Earl, Martin, McCarthy, & Soule, 2004; Karlsson & Sjøvaag, 2016). Recently, scholarly attention—especially of those working in political communication and social media—has shifted on the effects of the viral spread of disinformation and extremist alternative media over digital media ecosystem, prompting a debate about models, standards, and values of content and the ideational impact of political news and information. Grounded in the content that users share over social media, new categories and definitions have been brought forward (Kalogeropoulos, Negrodo, Picone, & Nielsen, 2017; Tandoc, Lim, & Ling, 2018). However, the current debate lacks a comparative framework of the types of information that circulate on social media that is grounded in methodologically collected data from social media sharing behavior as opposed to phenomenological evidence.

Certainly, propaganda, disinformation, and negative campaigning are not new phenomena of public life. But social media platforms offer new affordances for spreading political news and information at great speed and scale, while using an individual’s own data for targeted disinformation campaigns. Typologies in political communication have been useful for understanding the diversity of variables and cases, that both accurately describe the features of such new phenomena and identify transportable concepts across several cases, among real-world outcomes in which media accounts provide the primary features for important incidents (Althaus, Edy, & Phalen, 2001; Erickson & Howard, 2007). In particular, typology building allows for meaningful comparisons of the type and format of political news and information across countries (McMenamin, Flynn, O’Malley, & Rafter, 2013).

We advance this debate with a rigorously composed typology based on a focused, cross-country case comparison of social media content during three elections in large consolidated democracies in Europe—France, the United Kingdom, and Germany. These elections offered a unique opportunity for comparing political news and information across countries. The close proximity of campaign periods allows for a focused comparison that can both establish a content typology and set into sharp relief the differences that political culture might have in explaining what relative amounts of content are relevant in public life.

Methods and Sampling

To understand the circulation of political news and information during sensitive political events, we performed a real-time social media data analysis of political news and information shared on Twitter during three European elections in 2017. In all three countries, the elections generated large amounts of political debates and media coverage over Twitter. As a critical moment of political life, elections spark heightened public interest in political events which is also reflected in their voter behavior online (Howard, 2005). Boczkowski and Mitchelstein (2012) note that in times of heightened political activity, users shared more information on public issues.

Drawing from our real-time data collections, we developed a grounded typology of political news and information that is persistent across the three cases. Our methodology considers different sources of political news and information. This approach allows for a highly contextual evaluation which considers the reporting, design, and practices of content production as a whole. This comprehensive evaluation is especially necessary since junk news sources occasionally mix in factual reporting and news wire reports to signal legitimacy.

We collected data via public Twitter's Streaming API which comes with some known limitations. The Streaming API only collects at most 1% of all global public tweets related to a specific search query, but Twitter has not released the precise sampling method. The representativeness of its random samples has been questioned (Blank, 2017; Morstatter, Pfeffer, Liu, & Carley, 2013). Nonetheless, many scholars have recognized this approach to be highly useful for studying "ad-hoc publics" especially around political events (Burgess & Bruns, 2012; Larsson & Moe, 2012). Conducting the analysis in real-time has the advantage of capturing content before it can be removed or before non-permanent URLs become invalid.

Our methodology followed four stages. First, we manually identified relevant hashtags about politics and the election for each country and language. Usually these included the names of leading candidates, political parties, official campaign slogans, and salient election-specific issues. We pre-tested manually identified hashtags for relevance in test

data collections which we ran for at least a week in the month before the elections. From this dataset, we identified additional salient hashtags based on co-use and frequency of use. For each country, our team reviewed a list of 200 to 300 hashtags including their total number of shares for relevance. This sampling method allows us to capture salient political conversation about the elections which we are focusing on in this study. Our approach excluded minor hashtags that refer to small or short-lived issues. It does not capture conversation that users shared without using hashtags. A complete list of hashtags for each election can be found in Table 1.

Second, based on these hashtags, we collected tweets from Twitter that (1) contained the selected hashtags; (2) contained a URL to a web source, such as a news article, where the URL or the title of the web source included a selected hashtag; (3) retweets that contained a message's original text, wherein a selected hashtag is used either in the retweet or in the original tweet; and (4) quote tweets where the original text is not included but Twitter uses a URL to refer to the original tweet. Tweets with URLs that pointed toward another tweet were removed from our sample, as these tweets are generally generated automatically when someone quotes a tweet.

In total, we collected 843,146 tweets in France; 2,488,804 in the United Kingdom; and 984,713 in Germany. We then extracted all URLs that Twitter users in our sample shared. We identified 88,755 URLs in France; 724,988 in the United Kingdom; and 115,563 in Germany. If Twitter users shared more than one URL in their tweet, only the first URL was analyzed. For each of the countries, we prepared a random sample of 10% of tweets for coding. Coding decisions were then later applied to the entire datasets.

Third, we prepared our tweets for semi-automated analysis by providing a simple spreadsheet that listed the base URL to web sources, along with random samples of full URLs and frequency of shares. URLs that were shortened with a link shortener (such as bit.ly) were unwrapped with a script our team developed and added to the spreadsheet for analysis where possible. We also unwrapped URLs to social media posts that were sharing URLs (i.e., a tweet about a public Facebook post linking to a news story).

Finally, using an iterative coding process, our team developed a grounded typology and manually categorized base URLs, which is described in detail below. Our team labeled a total of 5,158 individual web sources of which 1,033 were shared in France; 2,579 in the United Kingdom; and 2,579 in Germany. Using a simple python script, the coding decisions for base URLs were applied to the individual links that user shared over the entire dataset. If sources were missing from our sample of 10%, they were categorized in this step. By using this method, we were able to classify and precisely classify 5,158 web sources. Following these consistent steps to sample, collect, clean, and code data allows for safer comparison and generalization across the country case studies.

Table 1. Sample Frames for Political News and Information Shared Over Twitter During Three General Elections in Europe, by Party Ideology, 2017.

Country	France	United Kingdom	Germany
Election date	23 April 2017, 7 May 2017	8 June 2017	24 September 2017
Sample date range	13 March 2017 to 19 March 2017	27 May 2017 to 2 June 2017	1 September 2017 to 10 September 2017
Tweets, total number	842,146	2,488,804	984,713
Tweets, mentioning candidate or party	620,373	1,147,123	620,280
Number of users	154,535	397,166	149,573
Tweets per users	5.4	6.3	6.6
Number of URLs	88,755	724,988	115,563

Source. Author's calculations based on data collected between March and September 2017.

Note. France's election proceeded with an initial round of voting in April and then a runoff election in May 2017. Sample collection for the French general election involved the following hashtags: French general election related hashtags included #présidentielle, #presidentielle, #présidentielles, #presidentielles, #presidentielle2017, #présidentielle2017, #presidentielles2017, #présidentielles2017, #electionpresidentielle, #électionpresidentielle, #electionpresidentielle2017, #electionspresidentielles2017, #elysee2017, #éllysée2017, #elysée2017, #jevot; Hamon hashtags included #Hamon, #Hamon2017, #BenoitHamon; Fillon hashtags included #Fillon, #Fillon2017, #FrancoisFillon, #FrançoisFillon; Le Pen hashtags included #LePen, #Marine2017, #MarineLePen, #MLP2017; Melenchon hashtags included #Melenchon, #Melenchon2017, #Mélénchon, #Mélénchon2017; Macron hashtags included #Macron, #Macron2017. Sample collection for the UK general election involved the following hashtags: UK general election related hashtags included general hashtags: #generalelection, #ge2017, #ge17, #generalelection2017, #election2017; Tory hashtags included #TheresaMay, #Tories, #Tory, #AskTheresaMay, #Conservatives, #VoteTory, #StrongandStable, #ToryManifesto, #ConservativeManifesto, #ImVotingTory; Labour hashtags included #VoteLabour, #Labour, #IC4PM, #Corbyn, #Labourdoorstep, #JeremyCorbyn, #Corbyn4pm, #JezzWeCan, #VoteCorbyn, #LabourManifesto, #ImVotingLabour, #ForTheMany, #VoteNHS; Liberal Democrat hashtags included #LibDems, #LibDemFightBack, #LibDem, #TimFarron, #UniteforEurope, #VoteLibDem, #LibDemSurge, #LibDemManifesto; UKIP hashtags included #UKIP, #Farage, #VoteUKIP, #Nuttall, #UKIPmanifesto; SNP hashtags included #ScotRef, #IndyRef2, #VoteSNP, #SNP, #NicolaSturgeon, #SNPmanifesto. German general election related hashtags include #btw2017, #bundestagswahl, #wahlkampf. AfD hashtags included #afd, #holderdeinlandzurück, #gauland. Christian Democratic Union hashtags include #angelamerkel, #fedidwgugl, #CDU. Free Democratic Party hashtags include #lindner, #denkenwirneu, #fdp. Bündnis90/Die Grünen hashtags included #grüne, #darumgrün, #diegruenen. Die Linke hashtags included #dielinke, #linke. Social Democratic Party hashtags included #martinschulz, #SPD, #zeitfürmartin.

Table 1 summarizes the comparative sampling frames for this study.

Typologizing Political News and Information on Social Media

The process of cataloging content involved several iterative stages, following best practices in both concept formation in typology building and content analysis (Collier, LaPorte, & Seawright, 2012; Earl et al., 2004). The crucial first stage of this process involves developing an initial training dataset, testing our assumptions and definitions on real data, which we then re-evaluated and refined over multiple rounds.

First, we developed a testing grounded typology for political news and information shared over social media with a test dataset. We began with a sample of URLs shared in the United States during that country's 2016 Presidential election—an important moment in which the “fake news” term first emerged. The team elaborated four broad: (1) professional news outlets, (2) established political actors, (3) polarizing and conspiratorial content, and (4) other sources of political news and information (Howard, Kollanyi, Bradshaw, & Neudert, 2017).

Within those categories, we developed a system of subtypes of sources, but instead of transporting a US-centric typology onto the political cultures in France, the United Kingdom, and Germany, our coding procedure often involved

full-team negotiations over subcategory labels, definitions, and coding decisions. This analytical flexibility allowed us to adapt as we met classification challenges, sometimes introducing new categories to account for differences in the three political cultures.

To examine junk news sources, coders browsed the domain for information about authors, owners, and funding sources, and corroborated this information with third-party sources such as *Wikipedia*. This helped inform the *Professionalism* and *Counterfeit* decisions made by the team. Coders also searched for information about the domain from other international and national third-party fact-checkers, such as *FullFact*, *Correctiv*, *AFP Fact Check*, and *Media Bias Fact*, to improve the certainty of criteria such as *credibility* and *bias*. *Style* was evaluated based on visual and aesthetic cues including images, wording, and videos used. Overall, this process allowed the team to develop a comprehensive typology of sources involved over 1,500 hr of coding, bi-weekly review meetings, and six training sessions. The typology reflects 22 months of iterative coding procedures, in which the typology was also exported to other country contexts in Mexico, the United States, Sweden, and Brazil (Glowacki et al., 2018; Hedman et al., 2018; Marchal, Neudert, Kollanyi, & Howard, 2018).

Subsequently, we developed a rigorous training system to train our team of language-specific experts in working with our typology. For this stage, we used a dedicated team of five

coders working on both US and European Union (EU) data in an iterative process. Coding and training on the shared English-language dataset from the United States achieved an intercoder reliability score of Krippendorff's $\alpha = .89$ among our groups of language-specific experts, signaling good concept formation across national contexts and high adeptness to our method. We chose the English-language dataset since all of our team members were fluent in the language which allowed for a comparative cross-country evaluation and meaningful assessment of the quality of coding decisions by country. To measure intercoder reliability, our team categorized web sources from our full list of base URLs in sets of random 50 sources. Individual cases in which the experts did not reach a consensus were re-evaluated in face-to-face team meetings. We did not calculate intercoder reliability scores within the language-specific teams of two to three coders, but as reported, the coders achieved sufficiently high scores.

Our system of analysis brought to light valuable distinctions across major categories of political news and information, and important nuances in subcategories. *Professional News and Information* was defined largely by the professional reputation of the organization behind the source, with two subcategories. *Major News Brands* included large sources that adhere to the standards and best practices of professional journalism, with known fact-checking operations and credible standards of production including clear information about real authors, editors, publishers, and owners. This content came from significant, branded news organizations, including any locally affiliated broadcasters. This content also included links to tabloids, unless they failed on at least three out of five criteria for junk news. This assessment is in line with recent political communication research, which finds that tabloids can exhibit "elective affinities" with disinformation and junk news (Chadwick, Vaccari, & O'Loughlin, 2018). *New Media and Start-Ups* were digitally native or start-ups that display evidence of organization, resources, and professionalized output that distinguishes between fact-checked news and commentary.

Professional Political Sources are produced by traditional political actors whose bulletins, working papers, websites, and reports provide citizen with information that is relevant in political processes. The subcategory *Political Parties or Candidates* was defined as official content produced by a political party or campaign. There were links to *Government* websites and reports from public agencies. Finally, content from *Experts* took the form of white papers, policy papers, or scholarship from researchers based at universities, think tanks, or other research organizations.

The third category of *Divisive and Conspiracy Sources* included various forms of unreliable and misleading reporting. The most important subcategory here was *Junk News and Information*. A source was labeled as junk news when it fulfilled at least three out of five criteria (professionalism, style, credibility, bias, and counterfeit). The subcategory *Russia*

consisted exclusively of state-funded Russian sources such as Russia Today and Sputnik to their known slanted reporting and application of propaganda techniques (Helmus, 2018).

The fourth category included a host of *Other Political News and Information* sources. These sources included many kinds of political actors who typically generate political news and information during an election. The subcategory *Citizen, Civic, and Civil Society* was used to label links to content produced by independent citizens, civic groups, or civil society organizations; watchdogs; and interest and lobby groups. It included blogs and websites dedicated to citizen journalism, citizen-generated petitions, personal activism, and other forms of civic engagement. We sparingly used a subcategory of *Other Political Content* to capture myriad other kinds of political content, including portals like AOL and Yahoo! that do not themselves have editorial policies or news content, survey providers, and political documentary movies. *Humor and Entertainment* was used to label sources that produce political jokes, sketch comedy, and political art.

Large numbers of URLs simply referred to content on other social media platforms, or were themselves spam, so we needed a major *Other* category. Links labeled as *Social Media Platforms* simply referred to other social media platforms, such as Facebook or Instagram. If the content linked to on social media platforms could be attributed to another source of information (e.g., when users were posting links to news articles posted on Facebook), it was. *Other Non-Political* were sites that did not appear to be providing information even though they were shared in tweets using election-related hashtags. Spam is also included in this category.

We also separated a *Language* subcategory for links that led to content that could not be interpreted by the coding team. An additional category *Not Available* included subcategories of links that were inaccessible after repeated attempts. Where possible, we wayback machine to code such sources.

There were three subcategories that were relevant in the training data but not in these three European cases. Links to unverified WikiLeaks content, political merchandise, and religious content were found in the US training data but rarely or not at all present in the case study data. If we found individual instances of such content, it was cataloged as *Other Political*.

Cross-Case Comparison

For our analysis of political news and information and automation, we collected data on Twitter for elections in France, the United Kingdom, and Germany. Tables 1 and 2 reveal details about the data collections including hashtags, dates, and candidates across countries. Twitter does not provide official information on active users by country, but according to the Reuters Digital News Report 2018, Twitter is the Top 2 social media platform in the United Kingdom and

Table 2. Political Communication on Twitter About Parties and Candidates During Three General Elections in Europe, by Party Ideology, 2017.

Ideology	France			United Kingdom			Germany		
	Candidate, Party	N	%	Candidate, Party	N	%	Candidate, Party	N	%
Liberal	Benoît Hamon, Parti Socialist*	99,898	16.1	Jeremy Corbyn, Labour Party	710,191	61.9	Christian Lindner, Freie Demokratische Partei	25,478	4.1
	Jean-Luc Mélenchon, France Insoumise	30,521	4.9				Katrin-Göring Eckhart, Cem Özdemir, Bündnis90/Die Grünen	15,705	2.5
Centrist	Emmanuel Macron, En Marche*†	267,842	43.2	Nicola Sturgeon, Scottish National Party*	137,931	12.0	Dietmar Bartsch, Sahara Wagenknecht, Die Linke*	14,751	2.4
				Tim Farron, Liberal Democrats*	63,027	5.5	Angela Merkel, Christlich Demokratische Union/Christlich-Soziale Union†	180,046	29.0
Conservative	François Fillon, Les Républicains	80,760	13.0	Theresa May, Conservative Party†	193,960	16.9	Martin Schulz, Sozialdemokratische Partei Deutschlands	87,642	14.1
	Marine Le Pen, Front National	141,352	22.8	Paul Nuttall, UK Independence Party	42,014	3.7	Alexander Gauland, Alice Weidel, Alternative für Deutschland*	296,658	47.8
Total		620,373	100.0		1,147,123	100.0		620,280	100.0

Source. This table involves a subset of the overall country samples detailed in Table 1. For France, authors' calculations from data sampled 13 March 2017 to 19 March 2017. For the United Kingdom, authors' calculations from data sampled 27 May 2017 to 2 June 2017. For Germany, authors' calculations from data sampled 01 September 2017 to 10 September 2017.

*The two candidates and parties with the largest portions of traffic—in that country's campaign period—generated by highly automated accounts.

†The winning candidate and party for that election.

Top 4 in France and Germany with growing—or in France stagnating—importance for news consumption (Newman et al., 2018).

In France, the first round of the French elections was held on 23 April 2017. Marine Le Pen of the conservative, nationalist party Front National and François Fillon from Les Républicains led in early polls (Lemarié & Goar, 2017). Emmanuel Macron ran for the central, socio-liberal En Marche! which he founded in April 2016. Benoît Hamon was the candidate of ruling the Parti Socialist. As no candidate won a majority in the first round, a run-off was held between the top two candidates Macron and Le Pen on 7 May 2017, which Macron won by a decisive margin. It was the first time since 2002 that a candidate from the Front National made it to the final round which some experts attributed this to a new populist movement after several terrorist attacks (McAuley, 2017). In the lead up to the French, several instances of far-right conspiracy theories and falsehoods with xenophobic, anti-Islam, and anti-Semitic references spread over social media (BBC News, 2017). A day before the second election, Macron was targeted with a fictitious email leak which was promoted by highly automated accounts under the hashtag #Macronleaks (Ferrara, 2017).

The UK election took place on 8 June 2017 and was announced just 2 months earlier by Prime Minister Theresa May. Theresa May ran for the Conservative Party. Jeremy Corbyn was Labour's candidate. The governing Conservative Party was confirmed as the single largest party in the House of Commons, yet lost its majority, which resulted in the formation of a minority government (Hunt, 2016). After the Brexit referendum in 2016, May had called for a snap election (Birrell, 2017). Targeted propaganda and the spread of misinformation had emerged as a critical concern on the public agenda in response to the Brexit referendum. In January 2017, the Culture, Media and Sports Committee launched an inquiry into “fake news” (House of Commons Select Committee, 2017). During Brexit, highly automated social media accounts were active for both the leave and remain campaign, generating large amounts of traffic (Howard & Kollanyi, 2016).

The German national election was held on 24 September 2017 and confirmed Angela Merkel in her mandate as chancellor for the fourth time. Angela Merkel led the Christlich Demokratische Union (CDU) and Christlich-Soziale Union (CSU). Martin Schulz was the Sozialdemokratische Partei Deutschlands (SPD) candidate. While Merkel's CDU won the vote, the party suffered substantial losses. After Schulz led the polls at distinct margin for months, the party lost the lead and achieved its lowest result in history (Frank, 2018). Represented in the Bundestag for the first time, the ultra-conservative AfD became the third party. Concerns over computational propaganda had emerged as a central political issue in the lead up to the election (Neudert, 2017). All of the major German parties committed to refrain using automation for campaigning, and German lawmakers introduced a stringent

law, the NetzDG, to counter hate speech and illegal misinformation (Stern, 2017).

Comparative Findings and Analysis

Candidate and Party Coverage on Social Media in Three Elections

Which political parties dominated Twitter conversation during their national elections? A number of political communication scholars have examined the central role of political parties for setting the agenda and influencing online discourse over social media (Howard, 2006; Kreiss, 2016; Nielsen & Vaccari, 2013). Our first point of analytical comparison shows patterns in how often each candidate and their respective parties were discussed on Twitter. This analysis reveals some of the salient quantities and qualities of the public conversation during the time of the national elections. Based on this evaluation, we can offer some basic context about which political parties were most relevant to the public conversation. Our method counted tweets with the selected hashtags in a straightforward manner. Each tweet was coded and counted if it contained one of the specific hashtags that were being followed. If a tweet contained more than one selected hashtag, it was credited to all the relevant hashtag categories.

Table 2 identifies the proportions of Twitter traffic. It distinguishes parties on a traditional liberal, centrist, and conservative ideological spectrum. Some of these parties are also considered populist—on the left or the right—or are essentially single issue, nationalist, or regionally specific parties. The craft of comparison involves selecting a parsimonious set of comparative attributes that take advantage of the cases in a limited case comparison, without creating points of contrast that have no real-world instances. But these three ideological categories provide the needed analytical contrast to make some rough associations across several political cultures.

In France, hashtags about Emmanuel Macron appeared most often—43.2% of the candidate-specific tweets during the week as a whole. For the most part, traffic about Marine Le Pen was about half that of Macron's. Yet, day to day, there were moments when Benoît Hamon drove the most traffic over Twitter, even though only 16.1% of the content overall related to him. In our follow-up of the run-off, Macron continued to dominate the Twittersphere with 42.3% of all traffic.

In the UK sample, hashtags about the Labour Party appeared most often, representing 61.9% of the party-specific tweets during the week as a whole. The Conservative Party generated a much lower proportion of the conversation at 16.9% and the Labour Party attracted consistently more conversation than the Conservative Party. At 12.0%, the Scottish National Party generated a disproportionately high percentage of the conversation, especially given the size of the party.

In Germany, Twitter conversations about the controversial AfD and its candidates were dominant on Twitter during

the German election and concerned 47.8% of all traffic. The traffic relating to the CDU accounted for 29.0% of all political traffic. Comparatively, traffic related to the Social Democratic Party and its candidate Martin Schulz accounted for only 14.1% of total Twitter traffic. There was a significant peak in traffic on 3 September 2017, the day of the public TV debate between Angela Merkel and Martin Schulz.

Table 2 provides yet another reminder of the distance between the volume of Twitter conversation about parties or candidates and their popularity among voters. Only in France did the ultimate victor in the election also dominate political communication on Twitter. In the United Kingdom, Jeremy Corbyn and the Labour Party dominated social media activity, while in Germany, the ultra-conservative AfD had the largest block of content. But neither of these parties won their elections.

Amplified Political Communication in Three Elections

How did amplification support the volume of political communication by particular parties over Twitter? We analyzed the activity of amplifier accounts over the three elections. We use the terminology of amplified traffic to refer to accounts that deliberately seek to increase the number of voices or attention paid to particular messages (McKelvey & Dubois, 2017). These accounts include automated, semi-automated, and highly active human curated accounts on social media.

We describe amplifier accounts as those that post 50 times a day or more on one of the selected hashtags during our sampling period. This detection methodology falls short of capturing amplifier accounts that are tweeting at lower frequencies. But, despite the simplicity of our metric, more complex methods using machine learning yield comparative numbers of false positives and remain contested in the field of computational social science. On the contrary, we have identified very few human users that tweet more than 49.5 times average per day on the sets of identified political hashtags. To reflect this observation, our simple threshold for identifying accounts that use some kind of amplification was chosen to reflect this threshold. Table 2 identifies the proportions of amplified traffic by candidate, party, and ideology. Automation trends are most sensibly compared between political, within countries, competing in the same election. An asterisk (*) identifies the two candidates or parties in each electoral contest that had the largest portions of traffic generated by highly automated accounts.

For the French election, it appears that a relatively low level of the overall traffic on the candidates was driven by amplifier accounts. At the high end, 11.4% of the Twitter traffic about Hamon was driven by highly automated accounts. At the low end, amplifiers generated 4.6% of the traffic about Mélenchon. Over the course of our sampling period, we find that the level of amplification remains fairly

Table 3. Political News and Information Shared Over Twitter During Three Elections in Europe, by Source Type, 2017.

Type of source	France		United Kingdom		Germany	
	N	%	N	%	N	%
Professional news sources						
Major News Brands	3,812	43.4	9,823	36.3	4,565	39.2
New Media and Start-Ups	286	3.3	3,385	12.5	114	1.0
Subtotal	4,098	46.7	13,208	48.8	4,679	40.2
Professional political sources						
Political Party or Candidate	1,189	13.5	2,030	7.5	1,047	9.0
Government	142	1.6	167	0.6	80	0.7
Experts	48	0.5	390	1.4	99	0.9
Subtotal	1,379	15.7	2,587	9.6	1,226	10.5
Divisive and conspiracy sources						
Junk News and Information	366	4.2	2,797	10.3	1,055	9.1
Russia	220	2.5	239	0.9	130	1.1
Subtotal	586	6.7	3,036	11	1,185	10.2
Other political news and information						
Citizen, Civic, and Civil Society	725	8.3	3,630	13.4	719	6.2
Other Political	352	4.0	1,223	4.5	995	8.5
Political Humor and Entertainment	55	0.6	959	3.5	378	3.2
Subtotal	1,132	12.9	5,812	21.5	2,092	18.0
Other						
Social Media Platforms	879	10.0	584	2.2	1,352	11.6
Other Non-Political	480	5.5	1,291	4.8	691	5.9
Language	115	1.3	261	1.0	127	1.1
Not Available	114	1.3	280	1.0	294	2.5
Subtotal	1,588	18.1	2,416	8.9	2,464	21.2
Total	8,783	100.0	27,059	100.0	11,646	100.0

consistent over the entire period, on average, 7.2%. During the run-offs between Le Pen and Macron, 16.4% of traffic was generated by amplifier accounts, with 14.0% and 19.5% of respective candidate traffic being driven by amplifiers.

It appears that in the United Kingdom, the parties and their respective candidates have a medium level of amplified traffic about them with fairly consistent levels across parties. On average, 16.5% of traffic about UK politics is generated by highly automated accounts that we are able to track. In our earlier analysis from the beginning of May 2017—soon after the elections were announced—we measured 12.3% of amplified traffic, so there was a rise in activity closer to elections.

The share of amplified traffic during the German election was not substantial. We identified 92 such accounts that generated 7.4% of the total traffic. The traffic generated by amplifier accounts for the CDU, FDP, and the SPD averaged between 7.3% and 9.4%. For the AfD-related hashtags, 15% of the traffic came from amplifier accounts.

Three points are notable across our cross-country evaluation. First, for all countries in our analysis, we find low to moderate levels of amplified traffic which suggests limited effects on aggregate social media sharing outcomes, overall in Europe. Second, the level of amplification in France and

the United Kingdom grew substantially closer to the elections. This suggests that amplification was coordinated to distort political debate surrounding the elections. Third, it is worth pointing out that we cannot establish who manages these accounts, and we do not analyze the content or emotional valence of particular tweets. Hence, this information alone is insufficient to determine whether the accounts are run by the campaign or candidate.

Sources of Political News and Information

What are the salient quantities and qualities of sources of political news and information users shared over social media? To understand what voters were sharing, we analyzed the web sources that users in France, the United Kingdom, and Germany shared links to. Table 3 breaks down the typological proportions of political news and information.

French Twitter users shared a lot of *Professional News Sources*, only occasionally shared content from major *Professional Political Sources*, and rarely relied on *Divisive and Conspiracy Sources* or other forms of political news and information. The largest proportion of content being shared by Twitter users interested in French politics—46.7% of all sources used—came from *Major and Minor News*

Brands. Information from *Political Parties or Candidates*, *Government*, and other *Experts* was somewhat less used. Yet, the largest proportion of that content is citizen-generated content.

The largest proportion of content being shared by Twitter users interested in UK politics comes from *Professional News Sources*, which account for 48.8% of all content. In comparison, *Junk News and Information* account for 10.3% of all content shared. A high percentage of other political content that was shared comes from *Citizen, Civic, and Civil Society*. *Russia* news sources, such as Russia Today or Sputnik, did not feature prominently in the entire sample.

In Germany, overall 40.2% of the political news and information Twitter users shared users discussing the German election came from *Professional News Sources*. Content from *Divisive and Conspiracy Source* accounted for 10.2% of shares. Links to content produced by *Government, Political Parties or Candidates*, or *Experts* altogether added up to just 10.5% of the total shares.

There are three main findings that should be noted across categories. First, our analysis demonstrates that in France, for every one link to *Junk News Sources*, there were seven links to content produced by professional news organizations. In the United Kingdom, the ratio was 5 to 1, and in Germany, the ratio was 4 to 1. The low proportion of junk news in France may be related to national political communication cultures which traditionally emphasize political debate in everyday life and social systems. Second, only a small number of domains were operated by known Russian sources during the time of the elections despite mainstream Russian outlets and their respective national offshoots having a substantial following in France, the United Kingdom, and Germany. Third, across national elections, the category of *Professional News Sources* was the largest category of sources shared during the elections that was at least roughly double as large as the second largest category of sources proportionally. Although many have suggested we are in “post-truth” area with declining levels of trust in the media, this finding suggests that social media users in Europe are still widely relying on traditional news media to inform their political decision making.

Conclusion

In each of these elections, social media had an important role both in providing an additional platform for public conversation and as a kind of news bulletin for political news and information about elections, candidates, and campaigns. Social media certainly has transformed the modes of political communication affording highly scalable, real-time debate and deliberation that is fueled by algorithms and personal data. In increasingly digital media ecosystems, the nature of political news and information has changed over time. Yet, current political communication theory lacks a comprehensive, grounded, and transportable typology of the

types of sources that social media users share as a fundamental framework for understanding the impact of such content on political systems.

In this article, we developed a grounded typology of sources of political news and information drawing from real-time social media data from three advanced democracies in Europe. Based on our comparative analysis, we put forward the concept of “junk news” which we use to refer to sources that publish or aggregate deceptive, conspiratorial, or incorrect information purporting to be real news and political information. This is the first multilingual, cross-national comparison of the production, consumption, and sourcing of political news and information over social media.

First, we find that political parties drive traffic in varying degrees. In France, the winning party led the public conversation, whereas in the United Kingdom and Germany, parties that challenged the current government, often controversially so, dominated traffic. Second, low to moderate levels of amplification intervene in political communication but activity levels rise closer to the elections. Third, in comparative terms, the importance of *Divisive and Conspiracy Sources* to political conversation varies from country to country with our analysis demonstrating that in France, for every one link to junk news, there were seven links to content produced by professional news organizations, and in the United Kingdom, the ratio was 5 to 1 and in Germany 4 to 1, respectively.

However, the most concerning trend for public life is not about the waxing and waning of professionally produced news as a source for political information. The real casualty appears to be political parties, government agencies, and experts. Trust in these sources of information has been in decline for a while, but in this comparative analysis, we are able to identify some sharp contrasts in how such sources are valued across political cultures in three advanced democracies. In France, for every one link to *Divisive and Conspiratorial Sources*, social media users shared more than 2.4 links to information coming directly from political parties, government agencies, or acknowledged experts. One might consider that ratio already too low. But in the United Kingdom, the ratio was the inverse, such that for each link to *Divisive and Conspiratorial Sources*, there was 0.8 links to sources from traditional political actors, and in Germany, the proportions were at par with a one to one ratio between these kinds of sources.

Assessing what information users shared on social media, and whether “fake news” even holds as a concept, was a project of immense analytical scale. We find “fake news” to be un-operationalizable and vague concept, but in analyzing data on what has been shared in recent elections, we offer an alternative typology that captures the poor, polarizing qualities of these sources of news and information. Grounded in what users were actually referencing during these three elections, we find that the substantive volume of political news and information does not even come from traditional sources of political and technical expertise. Rather than chasing a

granular definition of the popular “fake news” term, we focus our inquiry on an exhaustive analysis of various forms of political news and information and set it in a context of aggregate social media sharing patterns.

In comparison to our analyzed data from France, the United Kingdom, and Germany, French users were sharing much better quality information than what many US users shared, and slightly better quality news and information as German and UK users share. While we have yet to test this, we anticipate that extending the typology to countries in the Global South and non-democratic countries with less lively media ecosystems would require some substantial changes to the typology, such as introducing categories for pro-government state-funded media, and anti-regime blogs that are subject to suppression and censorship.

Many of the patterns in sourcing political news and information over social media—and the forms of content themselves—are new features of contemporary political communication. Researchers need to be flexible in applying the traditional definitions we have for what counts as political content, while interpreting the new types of sources and content. In this article, we develop and analyze a carefully constructed, holistic typology of how news production and consumption is evolving over social media.

Grounded typology building requires complex analytical exercise. But the political trends and events of the last few years paired with new modes of public discourse underscore that political communication scholars need to retool and adapt their frameworks. Typology building is a central exercise if researchers want to answer questions on the impact of problematic information on political processes, or critically interrogate how political discourse and news consumption are evolving over time. Most importantly, rigorous grounded typologies of content are fundamental for putting forward policy recommendations to inform platform design and content moderation in ways that underscore democratic ideals of free speech and deliberation.

Acknowledgements

The authors are grateful to Gillian Bolsover, Samantha Bradshaw, Clementine Desigaud, John Gallacher, and Monica Kaminska for their research assistance.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the European Research Council through the grant “Computational Propaganda: Investigating the Impact of Algorithms and Bots on Political Discourse in Europe,” Proposal 648311, 2015–2020, Philip N. Howard, Principal Investigator. For

supporting our Election Observatory and our research in Europe, we are grateful to the Adessium Foundation, Microsoft Research, and Omidyar Network. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funders.

ORCID iD

Philip Howard  <https://orcid.org/0000-0003-3380-821X>

References

- Althaus, S. L., Edy, J. A., & Phalen, P. F. (2001). Using substitutes for full-text news stories in content analysis: Which text is best? *American Journal of Political Science*, *45*, 707–723. doi:10.2307/2669247
- Aronovitch, H. (2012). Interpreting Weber’s ideal-types. *Philosophy of the Social Sciences*, *42*, 356–369. doi:10.1177/0048393111408779
- BBC News. (2017, April 5). Fake news in the French election. *BBC News*. Retrieved from <https://www.bbc.co.uk/news/world-europe-39495635>
- Birrell, I. (2017, May 2). Strengthen our hand in Europe? No, a landslide for May would weaken it Ian Birrell. *The Guardian*. Retrieved from <http://www.theguardian.com/commentisfree/2017/may/02/europe-landslide-victory-theresa-may-brexite>
- Blank, G. (2017). The digital divide among Twitter users and its implications for social research. *Social Science Computer Review*, *35*(6), 679–697.
- Boczkowski, P. J., & Mitchelstein, E. (2012). How users take advantage of different forms of interactivity on online news sites: Clicking, e-mailing, and commenting. *Human Communication Research*, *38*, 1–22. doi:10.1111/j.1468-2958.2011.01418.x
- Bradshaw, S., Neudert, L.-M., & Howard, P. N. (2018). *Government responses to malicious use of social media*. Retrieved from <https://www.stratcomcoe.org/government-responses-malicious-use-social-media>
- Burgess, J., & Bruns, A. (2012). Twitter archives and the challenges of “big social data” for media and communication research. *M/C Journal*, *15*(5), 1–7. Retrieved from <http://journal.media-culture.org.au/index.php/mcjournal/article/view/561>
- Chadwick, A., Vaccari, C., & O’Loughlin, B. (2018). Do tabloids poison the well of social media? Explaining democratically dysfunctional news sharing. *New Media & Society*, *20*, 4255–4274. doi:10.1177/1461444818769689
- Collier, D., LaPorte, J., & Seawright, J. (2012). Putting typologies to work: Concept formation, measurement, and analytic rigor. *Political Research Quarterly*, *65*, 217–232. doi:10.1177/1065912912437162
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 554–559. doi:10.1073/pnas.1517441113
- Earl, J., Martin, A., McCarthy, J. D., & Soule, S. A. (2004). The use of newspaper data in the study of collective action. *Annual Review of Sociology*, *30*, 65–80. doi:10.1146/annurev.soc.30.012703.110603
- Erickson, K., & Howard, P. N. (2007). A case of mistaken identity? News accounts of hacker, consumer, and organizational

- responsibility for compromised digital records. *Journal of Computer-Mediated Communication*, 12, 1229–1247. doi:10.1111/j.1083-6101.2007.00371.x
- European Commission. (2018, March). *A multi-dimensional approach to disinformation: Report of the independent high level group on fake news and online disinformation*. Luxembourg, UK: European Commission.
- Ferrara, E. (2017). Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday*, 22(8). doi:10.5210/fm.v22i8.8005. Retrieved from <http://firstmonday.org/ojs/index.php/fm/article/view/8005>
- Frank, A. (2018, February 10). Die kurze Ära Martin Schulz: Wacker gerackert [The short era of Martin Schulz: A fair bit of work]. *Spiegel Online*. Retrieved from <http://www.spiegel.de/politik/deutschland/martin-schulz-die-chronik-seines-scheiterns-als-chef-der-spd-a-1192809.html>
- Glowacki, M., Narayanan, V., Maynard, S., Hirsch, G., Kollanyi, B., Neudert, L. M., . . . Barash, V. (2018, June 29). *News and political information consumption in Mexico: Mapping the 2018 Mexican presidential election on Twitter and Facebook* (Data Memo No. 2018.2, Project on Computational Propaganda). Oxford, UK: University of Oxford.
- Hedman, F., Sivnert, F., Kollanyi, B., Narayanan, V., Neudert, L. M., & Howard, P. N. (2018, September 6). *News and political information consumption in Sweden: Mapping the 2018 Swedish general election on Twitter* (Data Memo No. 2018.3). Oxford, UK: University of Oxford.
- Helmus, T. C. (2018). *Russian social media influence: Understanding Russian propaganda in eastern Europe*. Santa Monica, CA: RAND Corporation.
- House of Commons Select Committee. (2017, January). *"Fake news" inquiry launched*. UK Parliament. Retrieved from <https://www.parliament.uk/business/committees/committees-a-z/commons-select/culture-media-and-sport-committee/news-parliament-2015/fake-news-launch-16-17/>
- Howard, P. N. (2005). *New media campaigns and the managed citizen*. New York, NY: Cambridge University Press.
- Howard, P. N. (2006). *New media campaigns and the managed citizen*. New York, NY: Cambridge University Press.
- Howard, P. N., & Hussain, M. M. (2013). *Democracy's fourth wave? Digital media and the Arab spring*. New York, NY: Oxford University Press.
- Howard, P. N., & Kollanyi, B. (2016). *Bots, #StrongerIn, and #Brexit: Computational propaganda during the UK-EU referendum* (Arxiv:1606.06356 [Physics]). Retrieved from <http://arxiv.org/abs/1606.06356>
- Howard, P. N., Kollanyi, B., Bradshaw, S., & Neudert, L.-M. (2017, September 28). *Social media, news and political information during the US election: Was polarizing content concentrated in swing states?* (Data Memo No. 2017.8). Oxford, UK: University of Oxford.
- Hunt, E. (2016, March 24). Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter. *The Guardian*. Retrieved from <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter>
- Kalogeropoulos, A., Negro, S., Picone, I., & Nielsen, R. K. (2017). Who shares and comments on news? A cross-national comparative analysis of online and social media participation. *Social Media+ Society*. Advance online publication. doi:10.1177/2056305117735754
- Karlsson, M., & Sjøvaag, H. (2016). Content analysis and online news. *Digital Journalism*, 4, 177–192. doi:10.1080/21670811.2015.1096619
- Kreiss, D. (2016). *Prototype politics: Technology-intensive campaigning and the data of democracy*. New York, NY: Oxford University Press.
- Larsson, A. O., & Moe, H. (2012). Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society*, 14, 729–747. doi:10.1177/1461444811422894
- Lemarié, A., & Goar, M. (2017, March 24). Affaibli, François Fillon joue le tout pour le tout [Affaibli, François Fillon puts all eggs in one basket]. *Le Monde.fr*. Retrieved from https://www.lemonde.fr/affaire-penelope-fillon/article/2017/03/24/affaibli-francois-fillon-joue-le-tout-pour-le-tout_5099944_5070021.html
- Machado, C., Kira, B., Narayanan, V., Kollanyi, B., & Howard, P. (2019, May 13–17). *A study of misinformation in Whatsapp groups with a focus on the Brazilian presidential elections*. Paper presented at the Proceedings of the 2019 World Wide Web Conference, San Francisco, CA. Retrieved from <https://dl.acm.org/citation.cfm?doid=3308560.3316738>
- Marchal, N., Kollanyi, B., Neudert, L.-M., & Howard, P. N. (2019). *Junk news during the EU parliamentary elections* (Project on Computational Propaganda, Oxford Internet Institute, Oxford University). Retrieved from <https://comprop.oii.ox.ac.uk/research/eu-elections-memo/>
- Marchal, N., Neudert, L.-M., Kollanyi, B., & Howard, P. N. (2018, November 1). *Polarization, partisanship and junk news consumption on social media during the 2018 US midterm elections* (Data Memo No. 2018.5). Retrieved from <https://comprop.oii.ox.ac.uk/research/midterms2018/>
- McAuley, J. (2017, May 8). Why the populists didn't win France's presidential election. *Washington Post*. Retrieved from https://www.washingtonpost.com/world/europe/why-the-populists-didnt-win-frances-presidential-elections/2017/05/08/061b5930-3023-11e7-a335-fa0ae1940305_story.html
- Mckelvey, F., & Dubois, E. (2017, June). *Computational propaganda in Canada: The use of political bots* (Working Paper No. 2017.6). Oxford, UK: Oxford Internet Institute, University of Oxford.
- McMenamin, I., Flynn, R., O'Malley, E., & Rafter, K. (2013). Commercialism and election framing: A content analysis of twelve newspapers in the 2011 Irish general election. *The International Journal of Press/Politics*, 18, 167–187. doi:10.1177/1940161212468031
- Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). *Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose* (Arxiv:1306.5204 [Physics]). Retrieved from <http://arxiv.org/abs/1306.5204>
- Neudert, L. M. (2017). Computational propaganda in Germany: A cautionary tale. In S. Woolley & P. Howard (Series Eds.), *Computational propaganda: Political parties, politicians, and political manipulation on social media* (p. 31). New York, NY: Oxford University Press. Retrieved from <http://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2017/06/Comprop-Germany.pdf>
- Neudert, L. M., & Marchal, N. (2019). *Polarisation and the use of technology in political campaigns and communication*. Retrieved from [http://www.europarl.europa.eu/RegData/etudes/STUD/2019/634414/EPRS_STU\(2019\)634414_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/STUD/2019/634414/EPRS_STU(2019)634414_EN.pdf)
- Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A. L., & Nielsen, R. K. (2018). *Reuters Institute digital news report*

2018. Reuters Institute. Retrieved from <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/digital-news-report-2018.pdf>
- Nielsen, R. K., & Vaccari, C. (2013). Do people “like” politicians on Facebook? Not really. large-scale direct candidate-to-voter online communication as an outlier phenomenon. *International Journal of Communication*, 7, 24. Retrieved from <http://ijoc.org/index.php/ijoc/article/view/1717>
- Persily, N. (2017). The 2016 U.S. election: Can democracy survive the Internet? *Journal of Democracy*, 28(2), 63–76. Retrieved from <https://www.journalofdemocracy.org/article/can-democracy-survive-the-internet>
- Stern, J. (2017, April 29). AfD verzichtet auf Meinungsroboter—oder nicht? [AfD not using social bots—or do they?]. Retrieved from <http://www.tagesschau.de/multimedia/video/video-283363~player.html>
- Swedberg, R. (2018). How to use Max Weber’s ideal type in sociological analysis. *Journal of Classical Sociology*, 18, 181–196. doi:10.1177/1468795X17743643
- Tandoc, E. C., Jr., Lim, Z. W., & Ling, R. (2018). Defining “fake news.” *Digital Journalism*, 6(2), 137–153. doi:10.1080/21670811.2017.1360143
- Tucker, J. A., Theocharis, Y., Roberts, M. E., & Barberá, P. (2017). From liberation to turmoil: Social media and democracy. *Journal of Democracy*, 28(4), 46–59.
- Vicario, M. D., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113, 554–559. doi:10.1073/pnas.1517441113
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359, 1146–1151. doi:10.1126/science.aap9559
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward and interdisciplinary framework for research and policy making*. Council of Europe. Retrieved from <https://rm.coe.int/information-disorder-report-november-2017/1680764666>
- Woolley, S., & Howard, P. (Series Eds.). (2017). Introduction: Computational propaganda worldwide. In S. Woolley & P. Howard (Eds.), *Computational propaganda: Political parties, politicians, and political manipulation on social media* (pp. 3–21). Oxford, UK: Oxford Internet Institute, University of Oxford.
- Woolley, S. C., & Howard, P. N. (2017, July). *Computational propaganda worldwide: Executive summary* (Working Paper No. 2017.11). Oxford, UK: Oxford Internet Institute.

Author Biographies

Lisa-Maria Neudert is a PhD candidate at the Oxford Internet Institute and a researcher at the Project on Computational Propaganda at the University of Oxford. Her research interest include social media and democracy and Internet policy in relation to countermeasures to computational propaganda.

Philip N. Howard is the Director of the Oxford Internet Institute and Professor of Sociology, Information and International Affairs. His research interests include computational propaganda, and the use of new information technologies for political action and social control.

Bence Kollanyi is a PhD candidate at the Corvinus University in Hungary and a researcher at the Project on Computational Propaganda at the University of Oxford. His research interests include the automation of social media accounts and the impact of bots on public life.