Three Consequences of Big Data on the Practices and Scholarships of Political

Communication

Fabio Giglietto

Università di Urbino Carlo Bo

Author Note

Fabio Giglietto, Università di Urbino Carlo Bo, fabio.giglietto@uniurb.it

Abstract

The advent of pervasive and connected digital technologies has profoundly affected both the way in which politicians interact with citizens and supporters and the way in which scholars study political communication. In the context of this ongoing transformation, this paper focuses on the role played by "big data" analytics. The large availability of digital footprints left by citizens in their everyday use of digital technologies have created possibilities for new forms of collaboration, advertisement, and propaganda. While political actors embrace these new strategies, media scholars studying political communication must adjust their methodologies in order to make sense of this mutated environment. This paper describes and discusses these three vectors of transformation in light of the new opportunities and challenges posed to scholars in the field of political communication.

Keywords: big data, micro-targeting, participation, attention, strategic amplification, data access

Introduction

The advent of ubiquitous connected digital technologies has profoundly affected the way in which politicians interact with citizens and supporters and, in turn, the way in which scholars

study political communication. In the context of this ongoing transformation, this paper focuses on the role played by "big data" analytics. The large availability of digital footprints left by citizens in their everyday use of digital technologies and social media has made possible new forms of collaboration, advertisement, and propaganda. While political actors embrace these new strategies, media scholars in the field of political communication must adjust their data gathering strategies and methods of data analysis in order to make sense of a profoundly mutated environment in which well-established approaches are not always reliable (Andersen, H. de Vreese, & Albæk, 2016), paving the way for conflicting empirical findings (e.g. the debate regarding the prevalence of echo chambers and filter bubbles).

In light of a discussion concerning the consequences of simplified massive coordination enabled by ubiquitous connected digital technologies on the boundaries of political movements and organizations, this paper describes two practices increasingly common in campaign strategies: social media-enabled micro-targeted advertising and new forms of propaganda and media manipulation. These three vectors of transformation are discussed in light of the new opportunities and challenges posed to scholars in the field of political communication. Each vector is discussed separately, and each is shown to have its own effect on both the practice and scholarship of political communication. Nevertheless, on a more general level, these consequences are deeply intertwined. The last paragraph discusses those consequences that do not uniquely pertain to a single vector.

Networked publics and the boundary of political organizations

Ubiquitous connected digital devices have enabled new forms of sharing, collaboration and collective/connective action (Bennett & Segerberg, 2012; Shirky, 2008). A wide range of online services allows users both to share their own content and re-share content created by other users.

Besides content, users also share – whether intentionally or not - their personal and behavioral data. This shared content and personal and behavioral data compose the raw material that enables a range of new forms of collaboration to arise. In essence, these collaborations are facilitated by a context in which group formation is streamlined and bonds between members are relatively weak (Granovetter, 1981). At the most basic level of this principle in action, a piece of content shared by a user can be aggregated to serve a collaborative purpose without the user having any intention of this themselves (e.g. your dog's public picture could be used on a website that showcases the main trait of a certain breed, or which trains machines in a certain learning algorithm). In this sense, the simple act of sharing something is potentially the starting point of a collaboration. On a more advanced level, a user may purposefully contribute their content or data toward a collaborative endeavor. Most of the time, in systems in which users are free to decide how much to contribute, a few users will contribute the majority of content while the rest is provided by a much larger number of occasional collaborators (Anderson, 2006). Such skewed distribution of labor has been observed both on Wikipedia (Park & Shin, 2014) and the community of developers of the Linux open source operating system (Benkler, 2011). Traditional organizations, including political parties, are formed around clear boundaries that draw a distinction between those inside (employers, affiliates, members) and outside. Communities

gathered around these new forms of networked collaboration are instead characterized by a blurred boundary between the outside and inside. While traditional organizations are ill equipped to accommodate occasional collaborations from outsiders, communities characterized by networked collaboration are, by definition, open to anyone's contribution. While most of the work is still performed by a small number of contributors, the limited contributions added by the large majority of members allows these networked communities to compete and sometimes even outperform traditional organizations. The loose boundaries between members enable these occasional forms of collaboration around temporary, shared goals to take place. Multiple

different users (and sometimes entire communities), each with their own individuality and agenda, can collaborate toward a shared, temporary objective without becoming part of a new stable entity with its own organization. Furthermore, where the shared goal is political, or pertains to a public issue, these forms of networked collaboration can become a vehicle for collective or connective action.

This kind of networked collaboration or collective/connective action is not new, but increasingly common. From the *Indignados* 15-M movement in Spain to the Occupy Wall Street protests in the United States, from the Yellow Vests in France to the *Umbrella* movement in Hong Kong, the actions performed by these networked forms of discontent have successfully raised media attention and repeatedly organized large crowds of protesters in the streets. The rise of digital parties (Gerbaudo, 2018) such as the Five Star Movement in Italy or Podemos in Spain can also be ascribed to this trend. Besides this phenomenon of protest movements morphing into official political parties, connective actions are also increasingly used by certain individuals and groups to back political parties while concealing any explicit links with those parties. By

employing this strategy, fringe groups that openly present themselves as white supremacist, misogynist or Islamophobic can simultaneously actively support mainstream parties/candidates. Given the lack of a clear connection, backed parties can deny having any link with these groups, their edgy operations, hate-speech, violence, and extremism.

The weak ties that characterize connective actions tend to blur the boundaries of traditional political organizations to the point that the analysis of official channels of communication only partially reflects the style, language and strategies employed. Making sense of the communication used by heterogeneous and leaderless movements that sprawl both official and unofficial blogs, forums, social media accounts and websites is increasingly challenging due to the inherent fragmentation of these movements. At the same time, focusing on the official

communicative channels of traditional political organizations and politicians only offers a partial view through which to understand the style, language, and implications of contemporary political communication when they are backed by these fringe, unofficially affiliated groups.

The implications of this networked organization also extend into traditionally well-regulated areas of political communication such as advertisement. The recent debate over whether to allow political ads on social media has highlighted the challenges faced when trying to define the boundary between ads regarding non-political and political/public issues. By also including ads concerning public issues, both Facebook and Twitter have recognized that banning ads run by verified political parties and politicians (e.g. candidates) is not enough.

Besides the challenges involved in drawing clear boundaries around political speech, to better understand the rationale behind banning or regulating political ads on social media it is useful to take a closer look at the second stream of transformations: online micro-targeted advertisements.

Micro-targeted advertisements

On March 17, 2018, a journalistic inquiry published by *The Guardian* exposed what has become known as the Cambridge Analytica data scandal. Exploiting a feature of Facebook's Open Graph launched in 2010 and then amended in 2013, the Cambridge researcher Aleksandr Kogan created a Facebook app called "thisisyourdigitallife". The app prompted users to answer questions designed to compile a picture of their psychological profile. Simultaneously, the app collected personal data from the profiles of those filling in the questionnaire, as well as their friends. By matching psychological traits determined by the questionnaire with the user's Facebook activities, Kogan built up a model designed to predict the psychological traits of any user based on their Facebook activity. Using this model, it was possible to target Facebook users with certain psychological profiles by exploiting the standard targeting function of the Facebook Ads platform. The data and model created by Aleksandr Kogan were then in turn exploited by

Cambridge Analytica and its clients, including Ted Cruz and Donald Trump, for aiming certain Facebook posts at specific micro-targeted sets of users who shared certain psychological traits.

While the benefits of this advertising strategy are questionable due to a general lack of compelling evidence of its potential to affect or shift political opinions (Broockman & Green, 2014; Hager, 2019), this example provides a good introduction to the idea of micro-targeted political advertising on social media. Micro-targeting, or the ability of an advertising platform to help an advertiser select exactly which users should see their ad, is not necessarily linked to psychological traits and can instead simply be based on user location, ethnicity, gender or interests (all of which are features made available out-of-the-box to the advertisers by most

social media platforms). For example, a voter suppression campaign may target citizens living in certain areas with the aim of limiting the turnout in regions where the opposition is stronger. The combination of such micro-targeted, promoted posts with the feature allowing Facebook page administrators to share certain posts with a selected subset of their audience ("dark posts"), enable politicians to address different audiences with different and potentially even contradictory messages. Much like the activist Stokely Carmichael, who used to shift the rhetorical style of his public speeches depending on whether he was addressing a black or white audience until he began addressing a broader public via television and radio in the 1960s (Meyrowitz, 1985), Facebook has enabled some politicians carefully to craft their messages for different forms of delivery to very specific audiences. While the combination of ads and dark posts are no longer allowed by Facebook, politicians continue to invest a growing share of their campaign budgets in online advertising.

To increase the transparency of political online ads, in October 2018 Facebook launched the Ad Library. At the time of writing, the Ad Library includes over 5 million ads about social issues, elections or politics created since May 2018 in the United States alone, and the total

amount spent on these comes to almost 900 million US dollars. In March 2019, the Library was expanded to include political and public issue ads from a total of 34 countries.

While this immense dataset of Facebook political ads clearly represents an incredible opportunity for scholars in the field of political communication, the use of the Library and its API (application programming interface, a communication protocol through which to query a dataset programmatically) for studying political communication still seems very limited, with only a handful of articles showing up on Google Scholar.

While this relative lack of research may simply reflect how recently the service was launched, some scholars and analysts have also raised questions regarding the shortcomings of the service (Bruns, 2019; Mozilla, 2019). One of the main issues raised has been in relation to the lack of information available concerning the targeted populations specified by advertisers. While the Ads Library shows both the region and age/gender/class profile of those who were shown the ad, it lacks detail concerning any additional targeting specifications (e.g. work, education or relationship status and interests). Due to this lack of details, it is hard to assess whether the observed regional and demographic distributions are in fact determined by hidden or confounding variables (e.g. older populations tend to be less educated, most young people are studying instead of working, etc.).

In this context, a heated public debate has arisen concerning the policies applied by social media companies to the advertising of public and political issues. Twitter's C.E.O. Jack Dorsey recently announced the decision to stop all political advertising on the platform globally (Conger, 2019; Twitter, 2019). Facebook, by contrast, was harshly criticized for distributing political ads without fact-checking them first (Vaidhyanathan, 2019). The intrinsic lack of transparency concerning the complex processes behind the formation of real-time aggregated targets from individual data prevents the development of evidence-based policies both within and outside of these platforms. At the same time, it fosters anxiety that, in turn, drives unsubstantiated concerns

regarding the potential effects of exposure to ads (Farkas & Schou, 2019). The Cambridge Analytica scandal both unveiled the naivety of social media platforms and the questionable intentions of individuals prepared to exploit these services (Mueller, 2019). However, in line with what has been observed by Yochai Benkler on disinformation, "evidence of action is not evidence of impact" (Benkler, 2019).

Unfortunately, evidence of the impact of exposure, such as proper causal relationships, are traditionally hard to obtain, even when data relating to the exposure to certain messages are available. This is not the case for online micro-targeted ads, nor for disinformation circulating both on- and offline.

Attention and amplification

While a vast amount of attention is devoted to micro-targeted ads, online advertisement is only the tip of the iceberg. Advertisement is only one of a wide and variegated set of strategies employed to drive eyeballs and clicks to certain online content. At the same time, it is not necessarily the most threatening or detrimental to the democratic process. Online advertisement is in fact tied to social media companies who may control the process and ultimately decide to suspend certain campaigns, or eventually the entire service. Online campaigns are also always financed by someone, making it easier to identify the actors behind certain operations by "following the money". Fringe strategies to drive attention are much more challenging to fight.

The role played by attention in contemporary media ecosystems is magnified by the abundance of information available. The multiplication of sources enabled by digital connected technologies has dramatically increased the competition and demand for attention. Attention thus become scarcer and its value skyrockets (Webster, 2014). In their analysis of the relationships between power and attention, Zhang and colleagues pointed out that attention is a necessary

condition for a speaker to change or mobilize the opinions of an audience. It is a channel through which an actor can ask others to take action; it also has a transferable value (those who receive

attention can also give attention - e.g. retweets) and has even become a signal of status in itself (Zhang, Wells, Wang, & Rohe, 2017). All social media platforms provide some sort of quantified attention metrics (views, likes, interactions, trending topics, etc.). While more detailed analytics are often accessible only to the channel/account owner, attention metrics are usually publicly available. A social media source or content that has proved successful in attracting attention is often seen as a valuable source or piece of content. Popular content tends to spread faster on social media due to algorithms that prioritize better-performing links, images, videos and posts. This measure of performance depends on an estimate of popularity based on the analysis of the quantified attention metrics provided by each platform. Besides the effect of this "the rich will get richer" feedback loop, popular social media content and highly discussed topics are often also featured in the traditional media, thus benefiting from a significant further spin (Phillips, 2018). The centrality of these metrics offers big rewards for those interested in increasing the visibility of certain content. This is not at all a new phenomenon. Fans' attempts to coordinate their behavior to push certain hashtags into Twitter trending topics date back at least to 2011 (boyd, 2017). During the last few years, similar practices have been increasingly observed with the aim of enhancing the spread of political news stories.

While the practice of influencing media coverage is not new, understanding the strategies of amplification (Donovan & boyd, 2019) in the era of the hybrid media system is crucial. Injecting and attempting to amplify the exposure of highly partisan and sometimes completely false narratives is a common strategy used by several fringe and subcultural internet groups (Marwick & Lewis, 2017). There is not a shortage of proof of the existence of these operations (Benkler, Faris, & Roberts, 2018; Bessi & Ferrara, 2016; Lewis, 2018). Nevertheless, it is harder than ever

to measure the prevalence of - and the effects of exposure to - these narratives.

The concept of exposure itself has also radically transformed. In a time of information scarcity, when relatively few media sources were available and people used to devote regular and specific time to getting news, it was relatively easy to ask about media consumption habits. However, media exposure today happens in a high-choice media environment, often occurs through channels that tend to display all sources equally, and news reach is more often driven by incidental encounters than by active selection.

The increased fragmentation of news media sources is well known (Fletcher & Nielsen, 2017; Messing & Westwood, 2014; Webster, 2014) but its consequences for the methodologies used by scholars in the field of political communication are often neglected. The growing and ever changing list of news sources to which citizens are incidentally exposed (Fletcher & Nielsen, 2018) hinders traditional questions based on self-reporting of news source consumption/exposure. Focusing on a list of mainstream media means excluding a large part of the audience. Providing a comprehensive and updated list is simply impossible. Once these lists are available (e.g. lists of debunked sources and problematic sources maintained by fact-checkers), it is then difficult to keep them updated, and it is risky to use outdated lists.

Furthermore, for the increasing share of the population exposed to news through social media, it can be hard to maintain a sense of which specific media brand's stories one has read due to the "all sources are equal" format of social media posts. This fixed display format weakens the role played by brand reputation. In this context, the concept of exposure itself become blurred.

Finally, the fact that a certain post is displayed in the feed of a certain user hardly says anything about the attention devoted to that post by that user. Moreover, data concerning views

are usually unavailable to researchers. In the feeds of social media platforms, links to political news stories compete within a broader range of different content. On the other hand, digital exposure can provide an alternative measure of attention by looking at the interactions around a certain post. Those who liked, retweeted, shared or commented on it were most probably actively exposed to the news story link in question or, at least, to its title displayed on the platform. Nevertheless, even by examining these interactions, one cannot be a hundred percent sure that the user ever visited the linked news story on the website of the original source. It is, in fact, common practice to interact with links on social media platforms without clicking through to and reading the related web page.

Discussion

While each of the analyzed vectors of transformation has its own repercussions on the practices and scholarship of political communication, further consequences result from the interactions of new forms of participation, micro-targeting ads, attention, and amplification.

A proper regulation of micro-targeted political ads, for instance, is hindered by the fuzzy boundaries of political organizations and rendered potentially ineffective by the alternative fringe strategies of attention amplification. Certain social media actors deliberately hide their identity (Donovan & Friedberg, 2019) or agenda and share highly partisan political content to unaware users who have subscribed to pages apparently dedicated to jokes or funny memes (Giglietto, Righetti, & Marino, 2019). Sometimes, elected politicians and candidates play an active, central role as the main hub of wide and loosely connected networks. In this role, politicians make use of their influence (in the online marketing sense) to amplify certain content (McIntire, Yourish, &

Buchanan, 2019) or to bring attention to certain users (Flegenheimer, 2019) in the form of

praise or attack.

The combination of increased accidental exposure to political news stories and the fuzzy boundaries around political actors fosters a context of permanent potential influence. While most citizens know how to resist persuasion attempts coming from traditional sources of influence (advertising, media and politicians) and in specific settings (e.g. while consuming news), it is much harder to deal with this in its new, ubiquitous forms. In line with what has been observed for personal influence by Lazarsfeld and colleagues in the seminal work "The People's Choice" (1944), ubiquitous persuasion attempts performed by unpredictable actors in an highly entertainment-dedicated space (e.g. social media) tend to reach undecided voters more frequently and catch their audience less prepared against influence.

References

Andersen, K., H. de Vreese, C., & Albæk, E. (2016). Measuring Media Diet in a High-Choice Environment - Testing the List-Frequency Technique. *Communication Methods and Measures, 10*(2-3), 81–98.

Anderson, C. (2006). *The Long Tail*. Random House Business Books (2006), Hardcover.

Benkler, Y. (2011). *The Penguin and the Leviathan: How Cooperation Triumphs over Self-Interest*. Crown Publishing Group.

Benkler, Y. (2019). *Cautionary Notes on Disinformation and the Origins of Distrust*. https://doi.org/10.35650/MD.2004.d.2019

Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda : manipulation, disinformation, and radicalization in American politics*.

Bennett, W. L., & Segerberg, A. (2012). THE LOGIC OF CONNECTIVE ACTION. *Information, Communication and Society, 15*(5), 739–768.

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday, 21*(11). https://doi.org/10.5210/fm.v21i11.7090

Broockman, D. E., & Green, D. P. (2014). Do Online Advertisements Increase Political Candidates' Name Recognition or Favorability? Evidence from Randomized Field Experiments. *Political Behavior, 36*(2), 263–289.

Bruns, A. (2019). After the "APIcalypse": social media platforms and their fight against critical scholarly research. *Information, Communication and Society, 22*(11), 1544–1566.

Conger, K. (2019, October 30). Twitter Will Ban All Political Ads, C.E.O. Jack Dorsey Says. *The New York Times*. Retrieved from https://www.nytimes.com/2019/10/30/technology/twitter-political-ads-ban.html

Donovan, J., & Boyd, D. (2019). Stop the Presses? Moving From Strategic Silence to Strategic Amplification in a Networked Media Ecosystem. *The American Behavioral Scientist*, 0002764219878229.

Donovan, J., & Friedberg, B. (2019). *Source Hacking: Media Manipulation in Practice*. Retrieved from Data&Society website: https://datasociety.net/output/source-hacking-media-manipulation-in-practice/

Farkas, J., & Schou, J. (2019). *Post-truth, Fake News and Democracy: Mapping the Politics of Falsehood*. Routledge.

Flegenheimer, M. (2019, November 2). What Happens When Ordinary People End Up in Trump's Tweets. *The New York Times*. Retrieved from https://www.nytimes.com/interactive/2019/11/02/us/politics/trump-twitter-retweets.html

Fletcher, R., & Nielsen, R. K. (2017). Are News Audiences Increasingly Fragmented? A Cross-National Comparative Analysis of Cross-Platform News Audience Fragmentation and Duplication. *The Journal of Communication, 67*(4), 476–498.

Fletcher, R., & Nielsen, R. K. (2018). Are people incidentally exposed to news on social media?

A comparative analysis. *New Media & Society, 20*(7), 2450–2468.

Gerbaudo, P. (2018). *The Digital Party: Political Organisation and Online Democracy* (1 edition). Pluto Press.

Giglietto, F., Righetti, N., & Marino, G. (2019). *Understanding Coordinated and Inauthentic Link Sharing Behavior on Facebook in the Run-up to 2018 General Election and 2019 European Election in Italy*. https://doi.org/10.31235/osf.io/3jteh

Granovetter, M. (1981). *The strength of weak ties : a network theory revisited*. Albany: State University of New York, Department of Sociology.

Hager, A. (2019). Do Online Ads Influence Vote Choice? *Political Communication, 36*(3), 376–393.

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). *The people's choice. 178*. Retrieved from https://psycnet.apa.org/fulltext/1945-02291-000.pdf

Lewis, R. (2018). *Alternative Influence*. Retrieved from Data&Society website: https://datasociety.net/output/alternative-influence/

Marwick, A., & Lewis, R. (2017). *Media Manipulation and Disinformation Online*. Retrieved from Data & Society website:

https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

McIntire, M., Yourish, K., & Buchanan, L. (2019, November 2). In Trump's Twitter Feed: Conspiracy-Mongers, Racists and Spies. *The New York Times*. Retrieved from https://www.nytimes.com/interactive/2019/11/02/us/politics/trump-twitter-disinformation.html

Messing, S., & Westwood, S. J. (2014). Selective Exposure in the Age of Social Media: Endorsements Trump Partisan Source Affiliation When Selecting News Online.

Communication Research, 41(8), 1042–1063.

Meyrowitz, J. (1985). *No Sense of Place: The Impact of Electronic Media on Social Behavior*. New York; Oxford: Oxford University Press.

Mozilla. (2019, April 29). Facebook's Ad Archive API is Inadequate – The Mozilla Blog. Retrieved November 15, 2019, from The Mozilla Blog website: https://blog.mozilla.org/blog/2019/04/29/facebooks-ad-archive-api-is-inadequate/

Mueller, Robert, S., & III. (2019). *Report On The Investigation Into Russian Interference In The 2016 Presidential Election* (United States. Department of Justice). Retrieved from https://www.hsdl.org/?abstract&did=

Park, H.-J., & Shin, K.-S. (2014). Pareto Ratio and Inequality Level of Knowledge Sharing in Virtual Knowledge Collaboration: Analysis of Behaviors on Wikipedia. *Journal of Intelligence and Information Systems, 20*(3), 19–43.

Shirky, C. (2008). *Here Comes Everybody: The Power of Organizing Without Organizations*. Penguin Press.

Twitter. (2019). Political Content. Retrieved November 24, 2019, from Twitter Business website: https://business.twitter.com/en/help/ads-policies/prohibited-content-policies/political-content.html

Vaidhyanathan, S. (2019, November 2). The Real Reason Facebook Won't Fact-Check Political Ads. *The New York Times*. Retrieved from https://www.nytimes.com/2019/11/02/opinion/facebook-zuckerberg-political-ads.html

Webster, J. G. (2014). *The Marketplace of Attention: How Audiences Take Shape in a Digital Age. How Audiences Take Shape in a Digital Age*. MIT Press.

Zhang, Y., Wells, C., Wang, S., & Rohe, K. (2017). Attention and amplification in the hybrid media system: The composition and activity of Donald Trump's Twitter following during the 2016 presidential election. *New Media & Society*, 1461444817744390.