Using social media data to study political science

Hossein Nourani1
Hitotsubashi University, Japan

Abstract

Social media are now involved in many aspects of human life. People use social media to do business, find friends, have fun, and discuss social and political issues. As an important aspect of our lives, and just like many other topics, politics is widely discussed on social media. Politicians, activists, civilians, companies and terrorists have made politics a hot topic on social media. However, there is no consensus among researchers on how reliable social media data can be for political science research, and what would be the proper method of collecting and analyzing social media data. In this paper, various theoretical and empirical works concerning the relationship between social media data, specifically Twitter data, and politics are critically examined to demonstrate how social media data affect politics and contribute to political research. The findings imply that social media data have significantly contributed to the field of political communication by offering inexpensive and easily-accessible information, empowering marginal social entities to participate in politics and internationalizing communication among political actors.

Keywords: social media, Twitter data, political communication, online social networks

1 E-mail: nouranih88@gmail.com

Received 2018/5/8, Accepted 2019/09/23
ISSN 2676-5535 (Online), DOI: 10.22067/ips.v1i1.72571
1. Introduction

Major social media platforms such as Twitter, Facebook, and YouTube have contributed to several political incidents around the world. Social media could be used by various opposition groups ranging from pro-democracy to terrorists. Belarussian activists used emails and weblogs to arrange their protest against the then-President Aleksandr Lukashenko in 2006. Later, Moldovan Protestors used Facebook, Twitter, and text messaging to spread their logic against the 2009 election results which would have kept the Party of Communists of the Republic of Moldova in power. Egyptian and Tunisian protesters managed to make use of Twitter to circulate their movements’ stories and mobilize potential supporters against the sitting presidents Mubarak and Ben Ali in 2011. On the other hand, the Islamic State of Iraq and Syria (ISIS) have resorted to Twitter, YouTube, and VideoCloud to express their violent ideology.

Even though the cheap and easily-accessible social media facilitate political activism, intelligence services do not hesitate to uncover social media algorithm and structure in order to track down opposition groups. For example, the Chinese government launched Sina Weibo microblogging website in 2009 to challenge Facebook and Twitter and exert effective control on domestic internet usage. Additionally, the US Department of State asked Twitter to postpone scheduled maintenance on its services to allow the flow of information to and from inside Iran when the 2009 Iran election results was widely challenged by domestic protesters (Pew Research Center, 2009).

Social media users have proven to be interesting citizen-reporters as they can produce a lot of live data about particular events that happens right in front of their eyes. Professional broadcasting companies have shown interest in using such loads of data to form their news programs, particularly when events such as wars or violent attacks are so intense that sending reporters to the field might not be the best idea.

This article examines several approaches towards the relationship between Twitter data and political science. Several important thinkers of this field are introduced along with the methods they created to collect and analyze social media data, particularly from Twitter to explain why this kind of data must be considered in political research. In this paper, first, social media is introduced. Second, the Twitter platform is briefly explained. Then, the importance of Twitter research is examined by bringing examples from the field of political communication. The challenges social media, and spe-
cially Twitter, pose to the field of political communication and their possible effects are also discussed. Finally, by referring to the examples brought up during the article, the importance of social media data in political research is examined.

2. Defining and conceptualizing “social media”

Social media services can be defined as Web 2.0 Internet-based applications through which individuals and groups attempt to create profiles. User-generated data is the main content of a social media service. Profiles and the data produced by users are in turn employed by each service to develop social networks (Obar & Wildman, 2015). Kaplan and Haenlein (2009) draw a line between the two concepts of Web 2.0 and User Generated Content before offering a straightforward definition of social media. Web 2.0, they argue, was first used in 2004 to describe the new way of using the Internet in which data and content were not produced by webmasters and administrators alone but all users would participate in modifying them. Accordingly, Web 2.0 was the basis for social media’s ever-growing popularity. On the other hand, User Generated Content refers to the data posted by social media users that is often innovative and accessible to a wide group of audience, though are not produced professionally. Social media are tools by which people produce and share information of different qualities outside the framework of traditional institutions. They are online facilitators of human networks that improve free social connectedness and thereby signify the democratic nature of the web (Farrington, Lee, Kilvington, Price, & Saeed, 2015).

Social media is also used by observers and researchers to refer to “online social networks” and “new media” depending on the level and perspective of analysis. When defining online social networks, Mislove (2009) pays attention to the importance of the availability of user profiles and the possibility of users’ access to other users’ profiles and content via existing links or creating new links in a way that they could maintain and strengthen their social ties. Cheung, Chiu and Lee (2010) define social networking websites as virtual communities that allow people to connect and interact with each other on a particular subject or to just “hang out” together online. In addition to descriptive personal profiles, members of such communities publicly articulate mutual “friendship” links with other members, creating a browseable network of social relations (Heer & Boyd, 2005). Chavonac and Dynel (2015) point to three possible types of involvement that social media grant:
recipient-to-recipient, with no involvement of the original author; recipient-to-producer, giving rise to across-the-fame interactions; and recipient-to-producer-to-recipient, with the recipient attaining the status of a (co-)author whose voice is re-mediated to new audiences by means of a “loop input” through the original producer.

Social media are particularly distinctive from the so-called “traditional media” or “mass media”, such as newspapers, magazines, radio and TV, in terms of technology, method of communication, content, use, as well as data production and consumption. The term new media, in its broad sense, refers to Internet-based media created using digital technology. The definition has evolved as various digital phenomena have gone online and provided Internet users with the opportunity to interact with each other, produce, and save data online. In addition to social networking websites, blogs and some collective video games offer their users online interactions. Such digital products as video games’ online services are usually free of charge, so all that users need to do is to connect their device to the Internet and open an account. Social media offer users unprecedented control over the message environment, allowing them to decide which messages to ignore and which to pay attention to (Buettner & Buettner, 2016; Hawthorne & Warner, 2015). Although social media and new media are similar in definition, we need to distinguish the two. In this research, the term social media simply refers to “Websites and applications that enable users to create and share content or to participate in social networking” (Oxford, 2016). Services such as Facebook, Twitter, and YouTube meet this definition.

3. Twitter features and functions

Twitter is a microblogging platform with 330 million monthly active users (Statista, 2017). Twitter messages or “tweets” are 280-character text-only messages that can include links, mentions, hashtags, and be liked and shared by users. “Retweeting” is sharing an original tweet with other users. Users often retweet tweets they endorse and agree with. So, retweets play a very important role in the flow of information throughout the platform even though they are unlikely to contain original content. Twitter users “follow” other Twitter accounts to keep up with those users’ activities such as their latest tweets and likes. Users can also “mention” other users’ usernames in their tweets. They often mention other users to react to their activities. Consequently, mentioning quite often signifies conversations among users. A
“hashtag” is a short phrase starting with a # sign which functions as “a hyperlink to the stream of recent tweets containing the specified tag” (Conover et al., 2011, p. 92). Twitter users make use of hashtags to search for Twitter contents including tweets, video, links, etc. that other users produce.

Despite the establishment of new services letting users share audio, video, and photographs, Twitter is not meant to host in-depth arguments and discussions. Twitter does not impose privacy restrictions or undertake content screening as much as, for example, Facebook. Twitter only investigates reports of abuse made by the victims or their authorized representative and does not screen content or remove potentially offensive or controversial content unless it breaches its ’Twitter rules’ (Farrington et al., 2015), whereas Facebook allows users to create secret groups, manipulate their page views, and report abusive contents. Twitter users might refuse to reveal part of their personal information such as job description or location, however, the main point of participating in Twitter discourse is to communicate with a wide audience. Ninety-five percent of the global Twitter users have set their account visibility to ’public’ (Bruns & Highfield, 2016). Twitter thus is potentially the most transparent social media platform because much of the discourse on Twitter is publicly accessible.

4. Twitter research affordances

Numerous studies have shown Twitter involvement in a wide range of political dynamics such as elections, revolutions, wars, and civil movements. Researchers are interested in data that can be extracted from tweets regarding specific events, and the way Twitter usage affects politics. In this section, I review a number of social media studies on political dynamics in order to explain by which methods and to what end researchers have investigated Twitter data.

Twitter research facilitates discovering socio-political networks, locating actors, tracking flows of information, making sense of various political discourses and behaviors. However, using Twitter data as a source of information has particular limits. Twitter records, as mentioned earlier, are not necessarily professionally produced so it is difficult to make sense of them. Therefore, one might need thousands or millions of Twitter records in order to form a satisfactory dataset for research. Besides, no matter how important and intense an event is, little Twitter data will be produced about it unless it is exposed to a high number of users a murder incident witnessed by a hun-
dred users would receive more Twitter attention than a massacre witnessed by no users. For example, one could find more Twitter data about a suicide in Berlin than a gang attack in Nigeria. Moreover, the more interesting users find a subject, the more data they will produce about it. Thus, complicated or secret issues would receive little Twitter coverage. It is also important to note that Twitter data does not last long; it is not possible to collect tweets from years, months, or even weeks ago. Furthermore, users tend to tweet on live and recent events rather than old ones. Access to Twitter data is not always satisfactory as the platform might block or remove certain accounts and data. For example, Twitter shut down the official Boko Haram account in January 2015.

5. Political Communication and Social Media

It is widely accepted that social media is a new sphere for political activity and participation. Habermas’ concepts of “structural transformation” and the “public sphere” have long been discussed in the field of political communication and media. The rise of social media is a challenge to both concepts. According to the theory of “structural transformation,” the public sphere is heavily mediatised by a few influential mass media agencies which form and direct political participation. In other words, public communication is operated by the elite who own or have access to mass media, and the public are less likely to influence the political arena. However, given the emergence of the Internet in the 1990s, news consumers have taken the opportunity to choose between news sources and challenge the supposedly media-controlled public sphere or, as Habermas puts it, the “mediated political communication” (Dang-Anh, Einspänner, & Thimm, 2014). The emergence of social media even seems to have challenged the mass media-dominated area of news production.

A new generation of critical theorists is testing Habermas’ analytical framework by analyzing the possibility of performing political action in social media. Some thinkers view social media as a new tool to reinforce democracy. For example, taking side with the advocates of critical theory and recalling the danger dominant media pose to democracy through occupying the public sphere, Winchester (2013) argues that the new media provide the critics of huge corporations with an opportunity to raise their voices and pressure the industrial world to care about issues such as global inequality and environment. Therefore, Winchester sees new media as a version of the
Habermasian public sphere where the civil society activist can effectively participate in making the world a better place.

Meanwhile, endeavors have been made to articulate new interpretations of political communication considering the latest evolutions in the field of social media. For instance, Chadwick, Dennis and Smith (2016) challenge the Habermasian approach to political communication that generally insists on the supremacy of mass media over the masses. They find a relationship in terms of democratic characteristics between this 20th-century-theory and what they call the ‘post-broadcast era’ in Western media systems. Chadwick et al. posit a hybrid media system encompassing the three fields of news and journalism, election campaigning, and engagement and mobilization. This hybrid media system is built upon interactions among older and newer media logics where the fields of social media and politics connect to each other. The hybrid media system is about actors’ judgment of “which medium or combination of media is most appropriate for shaping a political event or process” (Chadwick et al., p. 11).

As a case study, the article mainly refers to Obama’s 2008 Presidential campaign which made the most of online social networks for fundraising, image-making, capturing citizen input and mobilizing citizens for the campaign. However, it also notes that much of the campaign contents such as key slogans were first introduced in the mass media. Bringing up a few more examples, the writers claim social media is still only a secondary source for campaign information, and the hybrid media system does not always imply a more inclusive form of democracy, because elites also take advantage of this system. However, generally it is undeniable that the non-elite have been given opportunities to exert power in the hybrid system. Therefore, things are now more democratic compared to the condition explained by Habermas, decades ago.

Bruns and Highfield (2016) go further in criticizing the Habermasian theory through a multi-layer analysis of public behavior. They argue that the Habermasian concept of the public sphere has been challenged in the era of social media and no longer represents the contemporary global media ecology. They overview research on Twitter and Facebook to find alternative structures capable of replacing or augmenting the orthodox concept of the public sphere. The paper introduces alternative concepts of public spherules, ‘issue publics’, ‘mass self-publication’, and personal publics as subsets of the public sphere in order to make sense of fragmentations inside the so-called public sphere. Regarding social media, more specifically, the writ-
ers recall that social media activity usually starts with opening a personal account and getting in touch with a few ‘friends’ or ‘followers’ and thus creating a personal public to discuss private issues within a self-selected public. Users gradually consider the outcomes of their discussions and reports as a reliable source of information that partly shapes their larger concerns and ideas. Social media users then tend to look for a wider public with which they share their issues and concerns. At this point, the article implies that individuals are interested in taking part in debates related to a collection of specific themes they care about rather than meta-discourses. The writers then apply their conceptualization to the extended network of Australian Twitter users in order to produce an example of the current relationship between social media and publics. They conclude that leading mass media have lost their pre-digital era dominance, that people can choose between sources, and that local news markets have been disrupted due to the spread of the Internet. Social media have further mediatized daily life and allowed individuals to effectively and widely report and discuss their concerns online.

Wright, Graham, and Jackson, (2016) argue that social media is political and polarized, and argue for a comprehensive articulation of everyday political talk. They assume that the importance of everyday political talk lies in its role in forming public opinion and preparing citizens to take political action in the public sphere and the political system. The writers criticize the Habermasian model of deliberation for privileging reasoning in political communication and underestimating the informal nature of political talk. According to the article, everyday political talk takes place in a context and situation where informal talk and interaction are common, which is called the ‘third space’. Third spaces host non-political people who talk about political issues regardless of ideological borders in non-political communities. Social media, put Wright et al. (2016), are a recent version of this third space, but their limitations prevent them from meeting all of the third space’s characteristics. For example, the Twitter platform works well for broadcasting short news and views but conversations do not usually follow such short tweets. All in all, the article suggests that a comprehensive study of the flow of information across different platforms is required in order to understand how social media works as a third space.

Researchers are also still debating whether or not social media platforms either directly or indirectly shape their users’ posts and affect their socio-political ideas. Bunz (2016) follows Foucault’s take on the ‘truth’ which is
in turn inspired by the ancient Greek concept of ‘Parrhesia’, in order to discover the role of digital media in public debate. Parrhesia refers to telling everything about an issue despite danger and threat. Bunz (2016) argues that in the public sphere owned by a few, the marginalized try to form new publics and speak the truth. The new public, or the ‘counterpublic’, does not establish an antagonist relationship with the larger media but tries to reveal ‘the other side’ of stories. Based on this claim, Bunz (2016) proposes the term ‘Post-dialectical’ relationship to replace the antagonist relationship. Bunz argues that compared to traditional media, social media provide a better arena for the counterpublic to speak the truth because the content and their generators are much less controlled. Therefore, Bunz implies that the relationship between the traditional and social media is post-dialectical not antagonistic. Recalling the arrest of a Twitter activist for reporting a police brutality case in the United States, Bunz (2016) writes that Twitter is where the counterpublic speaks the truth despite danger. The paper concludes by claiming that social media has the potential to help the truth to be spoken, and if this is the case, it can also contribute to reporting various narratives of political events and thus take part in political dynamics.

Enli (2016) is not as optimistic as Bunz (2016) about finding the truth on social media platforms. Enli (2016) investigates authenticity illusions in political communication based on the assumption that authenticity is essential in social media politics. Enli (2016) studies two campaign videos in order to discover authenticity. The writer recalls that people gravitate towards social media because they have gradually lost complete trust in the mass media. However, candidates, as an example of political actors, are increasingly using media experts and professionals to facilitate their campaign with contents that look real, intimate, and cheap in order to create positive personal images of themselves. Enli (2016) investigates four elements—predictability, spontaneity, ordinariness, and imperfection—in two videos from Barack Obama and Jens Stolberg’s campaigns to discuss their techniques for presenting authentic images of themselves. Enli (2016) argues that the voters’ thirst for knowing about the candidates’ personal characteristics would lead them to what Enli (2016) calls an ‘authenticity illusion’ produced by the candidates’ media advisors. Enli(2016) expects such advisors to play crucial roles in politics in the future as personalizing of campaigns to attract voters will necessitate the manipulation of candidates’ images in social media.
In an empirical analysis, Barbera, Jost, Nagler, Tucker, and Bonneau (2015) mapped relationships between Twitter users according to their ideological tendencies. Barbera et al. (2015) studied ideological preferences of about 3.8 million users, and employed a data set including about 150 million tweets related to 12 political and nonpolitical issues to find out whether online communication is similar to an “echo chamber” or a “national conversation”. They try to estimate ideological stances of social media users. Assuming that people tend to contact others who have similar views, they mainly use the “follow” feature of Twitter to keep track of users’ ideological tendencies. Barbera et al. (2015) argue that estimating the ideological preferences of a large number of Twitter users greatly help researchers to investigate political communication as well as various forms of social and political behavior. They conclude that online communication is flexible and dynamic. In this sense, responses to different political and nonpolitical topics might either resemble an echo chamber or a national conversation. In some cases, public responses to a topic might fit both patterns in turn. The authors also claim that liberals were more likely to contact others who are dissimilar to them than conservatives do.

Díaz-Campo, Segado-Boj and Lloves-Sobrado (2015) apply content analysis to 688 tweets sent by ten Latin American leaders to find out the types of their messages and their purposes of tweeting in times of crisis. More specifically, they compare the number of tweets before and after a scandal, study the differences between the sitting presidents’ and major opposition leaders’ discourse, and analyze the politicians’ discourses towards the audience on Twitter. The time framework for analyzing each politician’s behavior consists of one seven-day period of a relevant crisis and the week from 23rd to 29th of September as a calm period of time in each of the five sample countries. The authors argue that the political leaders do not share a specific strategy for using twitter during challenging times, and that mass media still plays important roles in these events. These researchers code their data based on six variables: “tweet posts date”, “number of retweets”, “number of favorite tweets”, “mentioning of users in tweets”, “retweet or not”, and “author of the original message”. They then define the types of messages posted on Twitter in nine categories, namely, “information sharing”, “self-promotion”, “opinion/complaints”, “statements and random thoughts”, “me now”, “question to followers”, “presence maintenance”, “anecdote (me)”, and “anecdote (others)”. Díaz-Campo et al. (2015) employ these coding and categories to comparatively analyze the politicians’ ap-
proach in making use of Twitter, but fail to expose their common strategies in this respect. On the other hand, they claim that the rulers’ use of Twitter is mainly limited to constructing their public image, whereas the opposition figures aim to share their opinions and criticize the government.

Lotan et al. (2011) discuss information production and flows about 2011 uprisings in Egypt and Tunisia on Twitter as case studies in order to find out the role of Twitter in spreading news around the world. More specifically, the main theme of the research is to reveal the role of social media actors in disseminating information in Twitter in times of crisis when it is difficult to recognize the origins and validity of information. The writers believe understanding the dynamics of information flows is critical in reaching these objectives. As for data collection, API\(^1\) is used to query for the most recent Twitter posts every 5 minutes, requesting the last 100 publicly posted tweets containing specific words to create two sets of data: one for Tunisia and the other for Egypt. The researchers took four steps to identify an information flow including classifying tweets that were very similar to bins\(^2\), sorting bins by size (number of tweets included), choosing the top 10%, and then randomly choosing one-sixth of them to identify a total of 850 flows which are analyzed in more detail. Lotan et al. (2011) then classified actors in 12 types including involved individuals, mainstream media, and curious observers. The authors argue that the overlap between users in the two countries marks a pre-existing network of people who share political interests. People also may resort to Twitter to learn about trending topics. Even though individuals and unofficial networks rarely produce professional news, they are considered as more trustworthy than official broadcasting corporations in times of crises. Consequently, the most prominent retweet interactions happened between journalists and activists in both the Egypt and Tunisia datasets. The article recommends mainstream media to take advantage of the interaction between journalists and activists in order to reach to on-the-ground information. The researchers close the paper concluding they might have been somehow inaccurate because not only their datasets

\[^1\] Application Programming Interface is a set of functions and procedures that allow the creation of applications which access the features or data of an operating system, application, or other service.

\[^2\] In computational geometry, the bin is a data structure which allows efficient region queries. Each time a data point falls into a bin, the frequency of that bin is increased by one (Wikipedia, 2018).
do not contain all relevant datasets, but also the highly dynamic information hinder precise actor types classification. Moreover, the 10% sample might be biased.

6. Conclusion

Social media plays an important role in political life. Enormous amount of data about politics is produced on Twitter every moment. In fact, this is the latest version of flow of information among humans. Twitter analysis is important in political science because it reflects and affects political participation and communication in real life. The Twitter platform is a new arena for political communication, particularly for those who used to refrain from offline participation because of its cost and complexity. Politicians also take advantage of Twitter for image-making, vote-marketing, and mobilization. Thanks to social media, particularly Twitter, different narratives can be heard. Despite all of these facts, social media data is certainly worth studying in political science.

References


https://en.oxforddictionaries.com/definition/social_media


