

Social Media: The Site For Methodological Inquiry in Social Science Research

Murchana Kaushik, Queen's University, United Kingdom

The Barcelona Conference on Arts, Media & Culture 2021
Official Conference Proceedings

Abstract

This paper attempts to locate the significance of social media content analysis as a tool of data collection in social science research. In present times, it is seen that people increasingly use social media sites to share their opinions about issues of socio-political relevance. This paper analyses how recent studies on nationalism and identity have used public discourses on the social networking sites – Facebook and Twitter as tools for collecting data on the broad areas of identity formation, nationalistic assertion, etc. For this purpose, this paper analyses four studies that have highlighted the following four distinct instances of nationalistic assertions through social media platforms in different parts of the world - Donald Trump's campaign on Twitter for the 2016 Presidential elections, the importance of Facebook in the union flag protests in Northern Ireland in 2013, significance of Facebook in discourses on identity and ethnicity in Hongkong, and the use of Twitter in the discourse on refugee citizenship in Turkey. These studies have been purposively selected due to two primary factors- the novelty of their research approach and also because they are illustrative of four distinct cases of identity assertion representing four different cultures, societies, economies. By highlighting the advantages and the limitations of social media content analysis, this paper attempts to throw light on the significance of locating social media platforms as sites of methodological inquiry in social science research.

Keywords: Social Media, Methodological Tool, Social Science Research

iafor

The International Academic Forum
www.iafor.org

Introduction

“A typical user now spends 2 hours and 25 minutes on social media each day. Added together, the world’s social media users are expected to spend a total of 3.7 trillion hours on social media in 2021 – which would be ‘equivalent to more than 420 million years of combined human existence’” (Digital 2021 Global Overview Report, 2021). This metaphorical illustration appropriately highlights the magnanimity of the outreach of social media globally in present times.

The International Telecommunication Union (ITU) estimate states that in 2021, 63% of the total population of the world are internet users. According to the Digital 2021 Global Overview Report 2021, 53% of the world’s population use social media as of January, 2021. In the list of the most used social media platforms, Facebook, Youtube and Whatsapp occupy the top three positions, with each platform having 2000 million users or more. Amongst other popular sites are Facebook Messenger, Instagram, Tiktok, etc. The most common reasons behind people using social media are: News & Current affairs, entertainment, time-pass, to be updated about friends, keep others updated about oneself, research about products to be bought, social networking, share opinion, work networking, FOMO (Fear of missing out), sports, celebrity news, charitable causes (Digital 2021 Global Overview Report, 2021).

The available statistical evidence speaks volumes about the role and influence of social media in the daily lives of people all over the world. It is seen that people not only network with their close friends and family through social media platforms, but also rely on them for news, information about events, incidents, etc. Social media, therefore, has developed to be a fertile ground for observing and analysing social dynamics. This study tries to understand how social media has been used as a methodological tool to extract and analyse data by social science researchers in recent times. For this purpose, it takes into account four different studies which have used social media platforms as sites for data collection and analysis. The selected studies highlight four distinct instances of nationalistic assertions through social media platforms in different parts of the world - Donald Trump’s campaign on Twitter for the 2016 Presidential elections, the importance of Facebook in the union flag protests in Northern Ireland in 2013, significance of Facebook in discourses on identity and ethnicity in Hongkong, and the use of Twitter in the discourse on refugee citizenship in Turkey. These studies have been purposively selected because of the novelty of their research approach and also because they represent different cultures, societies, economies. I try to understand the methodology adopted by these studies, discuss the advantages and limitations of using social media as a research site, with a view to understanding a new methodological approach in the social sciences.

Social Media: The New Field

The entity of the ‘field’, its significance, and the ways to engage with it have been well deliberated upon by social scientists (Srinivas, Shah, & Ramaswamy, 2002; Wax & Wax, 1980; Davis, 1984; Everhart, 1984). The field serves as the location which enables a researcher to see, understand and analyse a phenomenon or event. It not only provides insights into aspects of social significance, but also facilitates the researchers to draw inferences and contribute to the knowledge repository. Malinowski (Malinowski, 1922), Srinivas (Srinivas, 1952), are examples of some early landmark field studies. The notion of the ‘field’ in these studies is however, that of a physical space which is out there, which is characterised by a visible social engagement among people, whereby the interplay of

institutions, values and norms in the daily lives of a society and its people can be seen. Contrary to this idea of a physical space and tangible activities, the social media platforms offer an alternative view of the entity of the 'field' for social science researchers. I call it an 'alternative view' because it does not possess the basic characteristics of the traditional field (physical space, real interactions, etc.), yet it is a space where social interactions and engagement take place and hence serves as a location for social science inquiry.

In this section, four distinct ways of conducting social science research inquiry will be discussed. With particular reference to four different studies conducted in different parts of the world, this section shall highlight various ways of using social media platforms as research tools. It is aimed that this account will not only reflect on the methods of using social media as a research tool, but also will bring to light the challenges and issues faced by the researchers in adopting social media as the field for data collection and analysis.

Facebook as a Site of Methodological Enquiry

A.

Li (Li, 2019) tries to understand how Facebook has been used as a platform for discussing and asserting issues of identity in Hong Kong. According to the 'Usage of Information Technology and the Internet by Hong Kong 2000 to 2019', as of 2019, approximately 94% of the households had access to the internet. It is estimated that approximately 85% of the total population of Hong Kong have access to social media as of 2021 (Digital 2021 Global Overview Report, 2021).

Li (Li, 2019) analyses how Facebook has been used as a medium by ordinary people of Hong Kong to discourse on issues of identity and nationality. Li (Li, 2019) identifies a fanpage called 'Hong Kong National History' (which is followed by over 5700 fans) and a digital publication called 'Hong Kong People's history of the thousand years' as the sources from which the required data for the study are collected. The mentioned page and its attached publication are run by a person called Eric who is identified as a Hong Kong nationalist. For the purpose of analysis, Li (Li, 2019) collects the posts from these sources for a distinct period of time (1 April to 31 December, 2017). In order to systematically arrange the data, the author classified the derived content into five different categories. The five categories are: quoted with/ without comment (they include posts shared by Eric from different sources), quoted news/ non news with/ without comment (this includes new/ non-news posts shared by Eric from different sources), self-created content (posts related to history or advocacy created by Eric himself), private post-post and other (they include posts originally created by Eric himself). Based on the schema developed and the time period defined, the author collected 427 posts from the fan page. Li (Li, 2019) then analyses the collected data by taking into account the responses that the collected posts have received from the fan page followers. It states that the highest number of responses (likes, shares and comments) are received by the posts that are originally created by Eric, and not the ones which he shares from other sources. Li (Li, 2019) states that Eric's popularity is apparent in the massive attention that he receives from not only the followers of his Facebook page, but also from the news media, public intellectuals, etc. Although Eric is not a political leader, nor a professional public opinion-maker, yet he is invited for media interviews, to deliver lectures and share his views, which are reflective of his popularity and acceptability amongst people.

Further, Li (Li, 2019) discusses some of the limitations of Eric's approach.

Further, Li (Li, 2019) also categorises Eric's discourse as civic studies or civic intellectuality, which is opposite to academic studies. Thus, on the one hand, Eric's posts may not be able to attract ordinary people due to the lengthy, historical, referenced nature of his work. At the same time, Eric's posts also run the risk of being attacked by academic scholars based on grounds of authenticity and quality of the content developed by Eric.

B.

Reillya & Trevisan (Reillya & Trevisan, 2016) analyse how Facebook played an important role in coordinating the first wave of the flag protests in Northern Ireland in January, 2013. It also discusses the ethical issues that may be involved in using Facebook as a site for conducting social science research. This study is based on data (16203 posts) collected and archived from the Facebook page called LPPU (Loyalist Peaceful Protest Updater) between the period of 2 and 22 January, 2013. The posts were collected and archived using the software called Discovertext. This particular period was used for data collection as this period is characterised by significant events in the Flag dispute issue. In order to systematically arrange and coordinate the data, this study followed the design of critical thematic analysis developed by Braun and Clarke (Braun and Clarke, 2013). The study, therefore, begins with the reading of the posts with a view to identify the key themes from the collected dataset. In order to derive themes from the Facebook deliberations, the study adopts the use of word clouds. This helped the researchers to identify key words, important themes which people engaged with in the Facebook discussions. The disadvantage of using word clouds is that it does not provide the context in which the words were used or discussed (Reillya & Trevisan, 2016). Thus, while it may be beneficial for researchers to the extent that it highlights important word or themes, at the same time it runs the risk of not contextualising the derived themes or word. Reillya & Trevisan (Reillya & Trevisan, 2016) further discuss some significant ethical concerns in using social media as a site for data collection.

One of the primary concerns for the authors was to decide on whether to engage in the process of data collection with the prior consent of the the Facebook page administrators, or to continue with data collection without informing the administrators. It may seem ethically sound to take prior permission or consent of the administrators. However, it runs the risk of making the administrators cautious of the content to be posted, which may further lead to manipulation of the contents.

Another ethical concern was how to present the collected data (posts) in a manner that it does not affect the reputation of the one who posted, while at the same time ensuring that the correct data/ information is reflected.

The other concern was whether to project the posts and comments of people in direct quotes. While on the one hand it would mean a transparent way of presenting data, on the other hand, it could also imply harm to the content creators. Thus, the authors decided to present the posts of only the known leaders of the flag protests in direct quotes, while abstained from doing that in case of the ordinary people (rank and file protestors) (Reillya & Trevisan, 2016).

Twitter as a Site of Methodological Enquiry

A.

Schertzer & Woods (Schertzer & Woods, 2021) study how Twitter was used as a medium of mass communication by Donald Trump during his 2016 Presidential campaign. According to the “Computer and the internet use in the United States: 2018” (2021), 85% of the households in the United States have broadband internet connection. Whereas, 82% of the total population use social media (Digital 2021 Global Overview Report, 2021). The study takes into consideration a total of 5,515 tweets which were sent by Trump during June 15, 2015 (Trump’s presidential campaign was announced on June 14, 2015) and January 20, 2017 (the day of Trump’s inauguration). All the tweets were collected from a public archive, which were then read by the authors. The tweets were then coded with the help of a software called NVivo. With a view to understanding the ethno-cultural content that Trump projected through Twitter, this study bases itself on a coded framework of five distinct categories of ethnic identity. The five categories include: People (It refers to the representation of an ethnic community through norms, values, beliefs, practices, etc.); Homeland (it refers to the representation of the community’s imaginary homeland); History (It refers to the representation of the ethnic community’s history); Religion (It refers to the representation of the ethnic community’s religion); Ethos (It deals with the idea of representing the ethnic community’s uniqueness). This framework has been adopted by the authors of the study from the work of Smith’s analysis (Smith, 1986).

B.

Bozdag (Bozdag, 2019) analyses social media representations of refugees in Turkey and discussed their role in shaping public opinion. According to the Digital 2021 Global Overview Report (2021), approximately 77% of the population are internet users and 70.8% of the population are social media users in Turkey as of January 2021. Bozdag (Bozdag, 2019) shows how Twitter can be located as a platform to understand people’s perceptions about the issue of refugees in Turkey. The study concentrates on the period between 2 July 2016 and 8 July 2016 because during this period Twitter saw a huge rise in the discussion on the issue of citizenship and refugees in Syria. The author searched for relevant tweets during the mentioned period with the keywords ‘citizenship’ and ‘Syrian’, and downloaded them with the help of a software called Dataminer. In order to systematically arrange the collected tweets, the author listed them in a MS Excel sheet. Out of the listed tweets, every fifth tweet from the list was lifted and listed on a new MS Excel sheet. These remaining tweets were used as the data that the study used for content analysis. Based on an analysis of these tweets, Bozdag (Bozdag, 2019) identified distinct positions taken by people on the issue of refugees and citizenship in Turkey. The study also tries to understand the meanings of tweets which may not otherwise be very clear due to the word limit defined by Twitter. To do so, the author tries to understand the contexts of such tweets by referring to newspapers, events in the local context, etc. The study further points out that one of the significant limitations of this method is that it runs the risk of including fake accounts in its data corpus (Bozdag, 2019).

Discussion and Conclusion

The above discussion brings to light some of the advantages and limitations of using social media as the site for social science research. Some of the prominent advantages of this method are:

- **Social media provides an easily accessible platform:** Social media provides an easily accessible platform for people to express themselves. It is seen that many people lack the confidence to express their opinions in the public sphere. However, despite the public nature of the social media platforms, people find it a convenient medium to express themselves, perhaps because of the virtual nature of engagement that social media offers. On the other hand, because of the public nature of the platforms, it is convenient for researchers to observe, collect and analyse required data from social media engagements.
- **Analysis through social media is a cost effective method:** From the researcher's perspective, data collection through social media is a non-expensive or a cost-effective way. The essential requirement in order to be able to use social media is an internet connection. Subscriptions to internet connections can be easily purchased at low costs in present times.
- **Analysis through social media saves time:** Data collection through social media is a less time consuming process as it is a computer-assisted process. Further, the use of softwares to collect, organise and store data saves the researcher the time and hassle of physically visiting the field and maintaining notes (in hard copies).
- **Social media provides updated information in no time:** It is seen that not only individuals, but also media houses, news publishers, organisations, etc are active on social media platforms. Thus, any event of significance immediately gets highlighted in social media through different sources. Therefore, this provides researchers with a fast access to the latest news, events, people's opinions on issues, etc.

The discussion of the four studies in the above section brings to light some prominent drawbacks of using social media as the site for social science research. The disadvantages limit the use of social media as a scientific tool of data collection and analysis. Some of the disadvantages are discussed below:

- **Quality of data:** Li (Li, 2019) in the study on identity issues in Hong Kong has highlighted that there is lack of any quality checking mechanism for the materials or information obtained from social media sites. As a result, the reliability and verifiability of the data obtained from social media are under question.
- **Authenticity of users:** Bozdag (Bozdag, 2019) in the study on social media representation of the issue of refugees and citizenship in Turkey states that identification of fake accounts is a major challenge. It is difficult to guarantee the authenticity of identity of the content creators on social media.
- **Maintaining the anonymity of users:** Reillya & Trevisan (Reillya & Trevisan, 2016) in their study on the role of social media in the Flag protests in Northern Ireland, state that they faced the dilemma of whether or not to reveal the identity of the social media users whose opinions (posts, comments, etc.) were considered for the study. While on the one hand, revealing the identity of the users would indicate the transparency adhered to by the researcher. On the other hand, revealing the identity of an user alongside his/ her opinions on an issue may turn to be harmful for him/ her.
- **Consent of users:** Reillya & Trevisan (Reillya & Trevisan, 2016) highlight another dilemma – Whether to take prior consent form the social media users before collection and analysis of data? While taking prior consent may seem to be an objective and ethical method, it however runs the risk of making the users alert and conscious of their activities on social media. As a result, their participation in the social media platforms may be manipulated and guided by concerns of reputational harm, of facing judgements, etc.
- **Representative sample:** Social media may not be a suitable platform for data analysis for researches which require not only a quantitatively sound sample size, but also demands a qualitatively representative sample size. For example, it is challenging to ensure that a

selected sample is representative of a balanced gender dynamics, class backgrounds, religious affiliations, national identities, etc.

- **Language variations:** It is seen that people often write in their native languages on social media. It may be challenging for researchers to obtain translations, understand, code and analyse the data.

The advantages and limitations of using social media as a site for social science research further bring out some important questions that may be of interest for further research. Some questions that research in this area should address in order to be able to make the optimal use of social media platforms are: How to ensure the security (anonymity) of the social media users?; How to identify a representative and authentic sample?; How to address the issue of linguistic variations/lack of translations that is commonly found in social media contents?

Acknowledgements

This study is a part of my PhD research on Intergenerational communication for social change in sectarian societies. I am grateful to my PhD supervisor, Dr. Gemma Carney, (Senior Lecturer in Social Policy and Ageing, School of Social Sciences, Education and Social Work, Queen's University Belfast), for encouraging me to work on this study. Her guidance and insights have helped me to explore the significance of social media for my PhD research and further develop this study.

I sincerely thank Queen's University Belfast for providing the financial aid required for me to register and present in this conference.

I am grateful to my friend Syed Ibtisam Tauhidi (Post Graduate Researcher in Computer Science, Queen's University Belfast) for his insights on various technical aspects associated with the usage of social media as a platform for data collection and analysis in social science research.

References

- Bozdag, C. (2019). Bottom-up nationalism and discrimination on social media: An analysis of the citizenship debate about refugees in Turkey. *European journal of cultural studies*, 23(5), 712-730.
- Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. London: Sage.
- Davis, J. (1984). Fieldwork. *RAIN*, 61, 2-3.
- “Digital 2021 Global Overview Report” by Datareportal:
<https://datareportal.com/reports/digital-2021-global-overview-report>
- Everhart, R. (1984). Dilemmas of fieldwork in policy research: A Critique. *Anthropology & Education Quarterly*, 15(3), 252-258.
- Li, S. (2019). Native history and nation building on personal online platform: Implications in Hong Kong context. *Social sciences*, 8,60.
- Malinowski, B. (1922). *Argonaouts of the western pacific*. George Routledge & Sons.
- Reillya, P., Trevisanb, F. (2016). Researching protest on Facebook: developing an ethical stance for the study of Northern Irish flag protest pages. *Information, communication & society* 19(3), 419–435.
- Schertzer, R. , Woods, E. (2021). Nationalism: the ethno-nationalist populism of Donald Trump’s Twitter communication. *Ethnic and Racial Studies*, 44:7, 1154-1173.
- Smith, A. D. (1986). *The Ethnic Origins of Nations*. Oxford: Wiley Blackwell.
- Srinivas, M. (1952). *Religion and society among the Coorgs of south India*. Oxford: Clarendon Press.
- Srinivas, M. N., Shah, A. M., & Ramaswamy, E. A. (2002). *The Fieldworker and the Field: Problems and Challenges in Sociological Investigation* (2nd ed.). Oxford University Press.
- “Usage of Information Technology and the Internet by Hong Kong 2000 to 2019” by Census and statistics department: The Government of the Hong Kong special administrative region:
<https://www.censtatd.gov.hk/en/EIndexbySubject.html?pcode=FA100109&scode=590>
- Wax, M., & Wax, R. (1980). Fieldwork and the Research Process. *Anthropology and Education Quarterly*, 11(1), 29-37.

Contact email: mkaushik01@qub.ac.uk
 murchana.kaushik@gmail.com