Chapter 37

Coding Twitter Data Using Qualitative and Computational Methods: A Mixed Methods Framework

Dhiraj Murthy

INTRODUCTION

Social media data have the capacity to give us insights into things that we have never been able to see before. Indeed, there is a highly diverse and vibrant set of literature just covering the popular microblogging platform Twitter. Big data methods have been successfully applied to a variety of contexts using Twitter data, such as examining differences in sentiment across population groups during the COVID-19 pandemic (Zhang et al., 2021), in relation to social movements (Abdul Reda et al., 2021; Barker-Plummer and Barker-Plummer, 2017), and elections (Alexandre et al., 2021; E. Chen et al., 2021). This work argues that many types of inferences about the social world can be made from Twitter data. However, one problem is that different media may facilitate particular types of communication (Yates and Orlikowski 1992). Twitter particularly affords in-the-moment content such as textual comments, photos, videos, links, and so forth. (see also McEwen and Fox, Chapter X, this volume). Though we may like to think that just about anything about human behavior can be deciphered from Twitter data, that simply is not true.

There are also challenges associated with data collection and analysis on Twitter (K. Chen et al., 2021), including the presence of bots (Samper-Escalante et al., 2021; Nevin et al., Chapter X, this volume), ethical issues in collection and sharing (Sloan et al., 2020; Webb et al., 2017; see also Jacobson and Gorea, Chapter X, this volume), sampling issues (Rafail, 2018; see also Quan-Haase et al., Chapter X, this volume), and difficulties in inferring demographic attributes such as age, race, and gender (Murthy et al., 2016). The focus of this chapter is Twitter, and how these specific challenges are associated with data collection and analysis on this microblogging platform. Several big data approaches that are popular for studying tweets are tremendously useful, but are often ill-suited to more in-depth contextualized analysis of tweets. This chapter therefore speaks to this debate and proposes the use of mixed methods approaches to create a more balanced means of analysis. For example, hand coding can be used to critically categorize tweets by addressing issues of ontology – our assumptions about the world. Specifically, coding categories can be emergent, undergoing several stages of reflection and engagement with theory in that domain (e.g. race, gender, and moral panics). What I mean here by 'ontology' builds from Hardt and Negri's (2005) argument of 'new ontology', which Murphy (2001: 22), succinctly defines as 'an innovative account of the being-in-process in which we are immersed'. Of course, we cannot reduce all subjective bias, but we can approach things like coding practice with some reflection on our ontological position. Hardt and Negri (2005: 312) argue that this type of critical 'new ontology' is part of their desire not to engage in 'repeating old rituals', rather, 'launching a new investigation in order to formulate a new science of society and politics …[that] is not about piling up statistics or mere sociological facts …[but] immersing ourselves in the movements of history and the anthropological transformations of subjectivity'. Descriptive logics, knowledge representation systems that 'subscribe to an object centered view of the world' (Baader, 2003: 351), rely on formal codification systems with strict notations and syntax. However, like all forms of classification, these are shaped by our worldview. In the case of the semantic web, for example, what metadata categories are deployed reflect a particular ontology, which can come from a privileged gender, racial, and/or socioeconomic position. Tweet codification systems are similarly affected, from what metadata is selected for study to how text, links, or hashtags are categorized.

Ultimately, this chapter presents an overview of means to categorize tweets, addressing issues of ontology and coding. It draws on qualitative approaches, such as grounded theory, to demonstrate the value of a solid coding scheme for the analysis of tweets. The chapter illustrates the value of this approach by providing examples taken from my own research of how this can be done in practice. Finally, the chapter draws some conclusions around a) the utility of mixed methods approaches for the study of Twitter as a platform, and b) the value of relying on established approaches, like grounded theory, to inform Twitter analysis. Emergent, open approaches to the study of Twitter-derived data not only advance what we can reliably infer from the popular medium, but also ultimately contribute to social knowledge. This chapter will first review our usual assumptions around Twitter research – especially around coding systems – before offering alternative approaches and operationalizing frameworks.

OUR USUAL ASSUMPTIONS

One assumption in studies using Twitter data is that the creation of knowledge from coded Twitter data is best served by closed coding systems, wherein attributes of tweet data (e.g. links, mentions, hashtags, and text) are given predefined coded categories. In these types of closed coding systems, categories are fixed and a research method(s) is/are applied. By contrast, an open system allows for codes to be altered or changed. Closed systems are common in the natural and medical sciences. In the case of the former, vegetation types, for example, have been classically categorized via closed coding systems (Ellenberg and Klötzli, 1972) and, in the case of the latter, as Stock et al. (1996: 109) observe, 'the common attributes of the variations of treatment conditions are listed and given code numbers'. In the context of Twitter, a preference for closed coding systems comes from computer science-based approaches. In Twitter's infancy, more interest tended to come from disciplines such as computer science and information systems (e.g. Bollen et al., 2009; Kwak et al., 2010) rather than from the social sciences (of course, exceptions exist, e.g. Hogan and Quan-Haase, 2010; Murthy, 2011b). Though this seminal work was pathbreaking, it led to a normative thinking that closed coding systems are better for studying Twitter data. A commonly held assumption behind the preference for closed coding systems was that computational approaches provided the best ways of studying Twitter data. This is evidenced by several studies, wherein frequencies of mentions or hashtags were used to summarize tweet data (e.g. Petersen and Gerken, 2021), rather than being part of a theory building exercise. The literature around Twitter as a detection system, wherein Twitter is used as a social sensor (Bornmann et al., 2020; Pandya et al., 2020) and physical sensor (Hernandez-Suarez et al., 2019) follows a similar logic of aggregating data and looking for trends and anomalies in tweets.

ALTERNATIVE EPISTEMOLOGIES AND ONTOLOGIES

Within the disciplines of computer science and information systems themselves, arguments were being made early on for hybrid or alternative methods to understanding Twitter data (i.e. Honeycutt and Herring, 2009). And when social scientists arrived after computer scientists to Twitter-related work, the call for critical epistemologies was renewed. This work has argued, for example, that hashtags and mentions imply complex social contingencies (Florini, 2014). Other work has argued that we need to be cumulative: tweet actions are accompanied by temporality, and a tweet at one time does not necessarily mean the same thing another time (Murthy, 2018). Additionally, digital ethnography has drawn from experience in ethnography and argues that learning more about the culture of a digital space is important (Murthy, 2011a; see also Fenton and Perry, Chapter X, this volume). In this sense, an emphasis is made on having experiential/cultural knowledge about a tweet corpus. This is viewed as integral to inquiry. Though it is not always apparent to Twitter researchers, Twitter is a ‘field’ in the Bourdieusian sense in that, as Lindgren and Lundström (2011) argue, the medium constitutes part of a social field with rules and presuppositions specific to it. Speaking from experience, I have often found myself deep in the field when coding a large sample of tweets. The process can be immersive, drawing one into a specific cultural context, as ethnography does for sociologists and anthropologists. It should be noted, however, that these types of ethnographic approaches are continually evolving to address the ubiquity of digital media today (Duggan, 2017) as well as shifts to app-based media cultures (Cousineau et al., 2019). Computational approaches have tended to shy away from these more ‘messy’ social scientific aspects of Twitter, which can also include contentious material (e.g. sexist, racist, homophobic content). However, ‘messy’, hard to code Twitter content (e.g. sarcasm) and users (e.g. transgender, multiracial, or transgeographical) have an important relationship to reflective inquiry.

METHODS

Twitter data can be very complex and are often poorly served by simply applying deductive reasoning. As boyd and Crawford (2012: 668) emphasize, ‘Big data is at its most effective when researchers take account of the complex methodological processes that underlie the analysis of that data’. This is not to say that traditional bottom-up inductive and top-down deductive methods are not useful for studying Twitter data. However, inductive and deductive methods have their own limitations and abductive methods, a form of reasoning ‘for finding the best explanations among a set of possible ones’ (Paul, 1993), were developed as an alternative approach out of responses to the reliance on model selection in the sciences (Bhaskar, 1976; Harré, 1976). In addition, mixed approaches to studying Twitter data open up possibilities.

In the case of Twitter work that is not altogether straightforward, other approaches can be highly beneficial. For example, with retroduction, a type of abductive method that emphasizes ‘asking why’ (Olsen, 2012: 215), researchers are able to probe the data regularly and to ‘avoid overgeneralization but searching for reasons and causes’ (Olsen, 2012: 216) instead. Or put another way, ‘the retroductive researcher, unlike the inductive researcher, has something to look for' (Blaikie, 2004: 972). In the context of Twitter research, retroduction emphasizes 'allowing for contradictory voices' (Olsen, 2012: x216. Similarly, Poole (2015) argues that retroduction allows us to be stopped by a surprise and then to try to comprehend it, enabling us to encounter problems and make sense of them. The idea behind such approaches is to highlight a sense of openness towards one's data and possible research questions.

CODING TWEETS

A variety of methods to code tweets and their users across a variety of languages, disciplines, and domains now exist (e.g. Ahn et al., 2020; Balakrishnan et al., 2020; Dann, 2010; Jha et al., 2020). Most of these frameworks examine the type of communication within tweets such as directed mentions via the @ symbol, replies, retweets, and general statements. Additionally, they also examine the types of content within tweets (e.g. promotion, personal reporting, link sharing) However, Twitter data generally follow a power law curve with a long tail – a distribution seen in many forms of empirical data (Clauset et al., 2009). This opens up interesting possibilities for selective coding. Specifically, a large percentage of mentions, hashtags, and links are directed to a small group of popular users, tags, and domains.

For example, Wu et al. (2011) in their 'Firehose' sample (a stream that includes all, rather than a sample of, posted tweets), found that a mere 0.05% of users they sampled accounted for almost half of the URLs in their collected data. Because of this distribution, it is possible to develop manageable coding rubrics and code small groups of users, domains, and hashtags, or other elements of Twitter data. This approach allows one to analyze large numbers of tweets (Wu et al., 2011) and can also be used to complement machine learning classification methods. Hand coding of tweets has also been used in a diverse range of contexts (e.g. Greenhow et al., 2021; Merkley et al., 2020) and is considered the gold standard. Given this, it is critical for researchers in this field to keep coding methods highly robust. However, automated methods using machine learning or other computational approaches continue to be used successfully (e.g. Hernandez et al., 2021) in certain applications (e.g. with less ambiguity in the tweet content or users being identified). Methods such as supervised machine learning use humans to manually annotate a sample of tweets and then use this knowledge to train machine-based systems for classification or annotation of content (e.g. Al-Laith et al., 2021).

Additionally, the types of communication the tweet indicates (e.g. promotion, referral, or personal status) can be coded. Robust ways to automatically classify whether the tweet contains a link, mention, or hashtag are readily available and can be combined with hand coding to discern more detail about tweet corpora. Additionally, coding methods can provide further detail on the types of users producing and consuming content in corpora and whether tweets were from an individual user, organization, bot, and so on. All of these standardized variables can either be human coded, machine learned, or some combination thereof (e.g. supervised learning). I have employed all of these with success (e.g. Ahn et al., 2020; Devaraj et al., 2020; Robertson et al., 2019; Xin et al., 2019). However, imposing preordained coding categories can limit our understanding not only of individual tweets, but also of larger Twitter discourses and the relationship between types of users and individual tweets. For example, the same text in a tweet could be serious when posted by an older user and sarcastic when posted by a younger user. Add race, gender, location, socioeconomic status, and a variety of other sociological variables and our ability to code with confidence can be significantly increased.

Although there are major pushes to move to exclusively computationally based coding models, there are major limitations to these approaches. Specifically, machines have difficulty coding nuance, interpreting hashtags/abbreviations/some slang, and deciphering visual content such as emoji, overlaid text, and GIFs. Mixed methods approaches can be particularly useful here. A larger argument I am making is that the ways in which we code social media data have enormous impacts on the empirical knowledge we are able to decipher from these data. Even if coding is systematic, it does not preclude miscoding. What I mean by this is that if coders are given coding rubrics that are leading, oriented around particular ontologies, or too narrowly defined, some content just gets missed. Of course, this can and does happen with interview-based coding. However, these coders usually have more context to draw from, given the much greater verbosity of an interview compared with a terse tweet. Brevity is not the sole factor here, as interviews are also often videotaped or audiotaped and gestural cues or intonation can assist with the success of the coding process.

GROUNDED THEORY

One tandem method that has been used in a variety of data-driven contexts including Twitter is grounded theory, a method that is premised on searching for possible explanations in the data rather than setting up hypotheses and testing them (an approach often ill-suited to Twitter-based research). The seminal work of Glaser and Strauss (2009) argued that reviewing collected data repeatedly and coding data into categories enables one to avoid some of the biases and limitations of overly positivistic research methods. Or, as Corbin and Strauss (2015: 22) highlight, 'the complexity of phenomena direct us to locate action in context, to look at action and interaction over time (process), and to examine action and interaction in routine as well as problematic situations in order to obtain a better understanding of how these relate'. Following Corbin and Strauss, my aim was not merely to code individual tweets, but to view tweets as part of a larger tweet 'context'. From this perspective, it is important to also understand the user who tweeted as well as the larger contexts they sit within. As Corbin and Strauss (2015: 7) discuss, a key feature of grounded theory is: '[T]he concepts out of which the theory is constructed are derived from data collected during the research process and not chosen prior to beginning the research. It is this feature that grounds the theory, and gives the methodology its name'. For this reason, employing emergent coding methods – though they are challenging – presents tremendous opportunities to understand tweets individually and collectively.

The advantage of this method is that it can also be combined with structured data. The ability to combine unstructured data such as status updates with structured data has utility for a wide variety of social media-related research work. For example, JavaScript Object Notation (JSON) data derived from Twitter provides structured fields such as 'user_mentions', 'hashtags' and 'in_reply_to_user_id_str'. Many social media application programming interfaces (APIs) deliver data via JSON. The format is useful for its readability by humans as it consists of a series of defined attributes and values rather than having abstract variable names or numerical variables. For example, an excerpt of selected fields from the JSON output for my Twitter ID is:

"user": {

"name": "dhirajmurthy",

"screen_name": "dhiraj murthy",

"url": "https://www.dhirajmurthy.com/about/",

"followers_count": 2063,

"friends_count": 849,

"statuses_count": 3381,

"time_zone": "Central Time (US & Canada)",

}

Figure 37.1 illustrates how I have holistically incorporated classic grounded theory approaches into the research design process. Specifically, the top of Figure 37.1 emphasizes that one should begin with the research problem and literature review, but not preordained research questions. Rather, as 'Field Research #1' in Figure 37.1 indicates, data should be collected and analyzed simultaneously, a process that leads to constant comparison and from there the generation of all possible conceptual categories or explanations.

With these larger conceptual categories in place (and their properties determined), one is able to then implement a coding method. Axial coding, where a category is placed in the center of the analysis and a set of relationships is created surrounding it, enables researchers to make connections between codes and to build explanatory models as part of a process of seeing relations between codes (Glaser and Strauss, 2009). This type of coding is iterative, as categories are placed and hypothetical relationships explored. For example, one could investigate situations causal to the category as well as the effects of the category and iterate as part of axial coding until the relationship sets are robust. However, there are limitations to this and 'Field research' adds value to the coding and analysis process. Specifically, 'Field Research #2' in Figure 37.1 emphasizes following some of the content embedded within tweets to contextualize the content. For example, after looking at linked URLs in a dataset, one may feel the need to code content at this point as this process could change what coding categories are deemed relevant. I provide examples in the next section regarding how this has unfolded in my own practice. 'Field Research #3' is a synthetic phase, where attempts are made to synthesize conceptual categories, refine their parameters, and prioritize core categories and theory. This phase provides an opportunity to reflect on how categories have developed. The core categories that have emerged can be part of a mixed methods project, where, for example, they are used with machine learning to extract tweets that might correspond to the categories of interest. Having this type of reflexive process also has the advantage of providing more nuanced categories that are then applied to computational big data methods including sentiment analysis.

[TS: Insert figure 37.1 here]

IN PRACTICE: A CASE STUDY OF #ACCIDENTALRACIST

Operationalizing these types of frameworks does require a different ontology of tweets in the sense that many of our approaches to studying Twitter are often closed. This may come as a surprise to some. Computational methods have also become firmly established in the literature and systematic reviews confirm the pervasiveness of machine learning methods in studying tweets (Kumar and Jaiswal, 2020). Because computer science and information science have historically been the majority producers of research that uses Twitter data (Zimmer and Proferes, 2014: 252), the dominant ontological worldviews in these fields have had great influence on how we study Twitter data. Mixed method approaches such as described in Figure 37.1 require one to be open in the inquiry, allowing coding to be emergent. Tweets are not merely bits of text. We, as researchers, have a real opportunity to ask what is happening in the tweet and to think about Twitter API-derived JSON data holistically. For example, Manovich's (2001) notion of 'digital objects' can be useful in thinking of tweets as a complex entity. Specifically, in the context of web pages, tweets can be thought of as 'interfaces to a multimedia database' (Manovich 2001: 37). In addition, as Quan-Haase et al. (2015) argue, the context of social media use matters (in their case, communities of digital humanities scholars shaped the content and organization of tweets). Just like any text can be taken out of context, so can tweets.

A key aspect of this is to think openly about what collected tweet data are helping us to study, broadly speaking. Again, this requires a certain ontological openness to the research process. Corollaries to this are:

a. Are we being reflexive on the point of view/standpoint we are interpreting? and
b. Are we being flexible or following prescribed rules?

Though these steps can be seen as a barrier to the Twitter research process, I firmly believe they open up exciting new lines of research possibilities. For example, Figure

37.2 visually illustrates how I adapted Corbin and Strauss' (2015) model specifically to Twitter data.

[TS: Insert figure 37.2 here]

This model leverages a continual collection and analysis method in order to discern social knowledge that is not straightforward. After raw tweet data are obtained (the 'COLLECT' phase), the API may need to be queried regularly in the 'CONTINUED' phase to study relevant conversations, images, followers, other hashtags, external media, and so forth. Figure 37.3 illustrates how I applied this to work on #accidentalracist, a hashtag associated with a controversial duet by Brad Paisley and LL Cool J that received significant attention on Twitter and became a trending topic (Muse, 2013). The hashtag covered comments about the song, which mixes country and rap, race (as the song refers to slavery, the Confederate flag, and the 'white hoods' of the Ku Klux Klan), and various interview gaffes by the artists. These data presented a very complex social engagement with the album that ranges from dismissive to supportive as well as involving various levels of richness. In other words, there is a discursive value to the hashtag. However, as is common with Twitter data, there is also a lot of noise and this presents difficulties in terms of discerning the relevance (i.e., 'signal') in a set of tweets (O'Neal et al., 2018).

[TS: insert figure 37.3 here]

Figure 37.3 illustrates how I overcame some of these issues by an iterative process of coding and analyzing data, which, like many qualitative methods more broadly, views this as a journey that does not 'follow a straight line' (Bryman and

Burgess, 2002: 208). For example, I went back and expanded URLs and added top-level domains to my dataset. I also followed some of the top links that revealed important media sources, such as an article by *Essence Magazine* (2013) that briefly described the album and asked its reader base of African–American women to vote whether the song helped race relations or not. Operationalizing this type of ontology requires several stages of coding. Key to this approach is to be open to diverse messages in one's data, as the example in Figure 37.3 illustrates.

Figure 37.4 illustrates how memo making during collection and analysis is a 'crucial step' (Charmaz and Mitchell, 2001: 167) to both coding development and theory building. Also, comparisons across diverse data at each stage provide reflexivity and triangulation, rather than proving particular paradigms. As Figure 37.4 illustrates, memo making at each stage is integral as this allows researchers to be open to what knowledge Twitter data can help build. In this framework, I began sampling by date (age, hashtag, seed user, Twitter list, etc.) and completed the preliminary '1st stage coding'. I then actively collected further data like linked images, while coding. Specifically, this coding process guided what further data I needed. (I am currently deploying this same framework on a study of Parler posts in which we have sampled 200 posts ('parleys') and are iteratively trying to make sense of their content.)

I iteratively compared codes and subcodes with JSON-attributed data that I was receiving from Twitter calls. Specifically, there were patterns and themes emerging in other JSON attributes (e.g. in the language code or whether a user was 'verified') that affirmed or challenged my previously established codes. This is a juxtaposition to merely looking at queries run against a comma-separated values (CSV) file or even the CSV file itself. As part of this process, I used memo making of JSON responses I got from Twitter as I did not actually use all the fields in my CSV file. In this framework, the 'who', 'what', 'when', 'where', and 'why' are all kept open to interpretation in the coding of content. In other words, the larger dataset with full JSON data delivered by the Twitter API is kept as a resource during the grounded theory process, wherein the subset of filtered data by specified variables could be augmented with other variable fields during the research process if a value to doing this arises. The idea here is to navigate these data in different ways and to see what coding categories are determined. Emergent patterns can be captured well in this method. The traditional approach is to apply preordained coding rubrics for tweet data. However, if tweets are treated as 'digital objects' (Manovich, 2001) open to nuanced forms of interpretation, we can have richer understandings of tweet corpora (although we do have to deal with smaller *n* counts). First, this method enabled me to gain new insights into #accidentalracist content. Specifically, an iterative approach kept me looking for more and led to me encounter differing viewpoints amongst African–Americans about the song. Second, this approach enabled me to dig deeper than a binary classification (racist/not racist), which strictly closed, computational approaches favor. By doing so, I found that the articulation of race and racism in #accidentalracist tweets was highly nuanced. It was not just white users supporting the song and Black users condemning it. Rather than ending up with blanket generalizations that paper over the cracks, I found examples of thoughtful interrogations specific to a user's gender, location (e.g. urban versus rural responses), and even educational background (e.g. some claimed educated Black people had been 'whitewashed' and would probably like the song).

Indeed, different things caught my attention (like language code) that popped up in my JSON observations. Again, the idea here is to see what surprises you. These surprises have helped me to drive ‘2nd stage coding’ into specific areas. Indeed, even the theories I have engaged with have evolved greatly during the research process. This is in contrast to deciding on a fixed theory/theories to test the research questions. This has been particularly true in my disaster-related (Murthy and Gross, 2017) and race-related (Murthy and Sharma, 2019) Twitter work. Additionally, I have noticed the role of humor in particular corpora by using this approach, where humor came into 2nd or 3rd stage coding, as I found in the case of tweets posted during Hurricane Sandy (Murthy and Gross, 2017).

[TS: Insert figure 37.4 here]

USING COMPUTATIONAL APPROACHES TO PROBE TWITTER DATA

My focus so far has been on collecting data and using qualitative and mixed approaches in the first instance of Twitter analysis. However, computational approaches can also come first, yielding data that can be incorporated as part of ‘Field Research #1’ in Figure 37.1. Human, manual coding can then occur (following Figure 37.1) and this coding can be informed by machine learning techniques applied to tweet content, profiles, and other metadata. Such methods can also advance computational approaches (e.g. via supervised learning).

A method I have explored many times across a wide variety of social media including Twitter data is latent Dirichlet allocation (LDA) (see also Huner and Suárez, Chapter X, this volume), a Bayesian ‘topic model’ approach that uses computational machine learning methods to derive topic clusters. LDA works by reading in text and then a discrete number of topics are generated (generally not more than 100). LDA 'is a robust and versatile unsupervised topic modeling technique, originally developed to identify latent topics …[with a] probability distribution over words (as opposed to a strict list of words that are included in or excluded from the topic)' (Gross and Murthy, 2014: 39). These topics are sometimes straightforward and at other times indicate unexpected or surprising interactions. Figure 37.5 illustrates three LDA-derived topic clusters from a 50-topic LDA application to 90,986 cancer-related tweets including the following keywords: cancer, mammogram, lymphoma, melanoma, and cancer survivor. As Topic 5 illustrates, one topic of collected cancer-related tweets refers to family, friends, hospitals and indicates a topic cluster around procedures/diagnoses. Topic 7 (which is only partially listed due to space constraints) starts with a diverse array of words, but then moves to beauty and later on down the list are words like makeup and lipstick. I would not have set out to understand these sometimes peripheral aspects of tweeting and cancer, but it turned out 'looking good' and keeping up beauty rituals was very important for a significant number of Twitter users. Another topic (not included in Figure 37.5) indicated subjects surrounding cancer and pets, which I discovered often involved lymphoma in dogs.

[TS: Insert figure 37.5 here]

Ultimately, one can effectively use machine learning approaches such as LDA to derive topic clusters around a Twitter corpus. Figure 37.5 illustrates this application with cancer-related tweets, but I have similarly used LDA and other machine learning methods on a variety of Twitter and other social media corpora (Alexandre et al., 2021; Xin et al., 2019). Another opportunity for coding arises here. This section emphasizes that it is possible to also have the computational element come first. Specifically, I have effectively used machine learning approaches such as LDA to derive topic clusters around a particular Twitter corpus. I have then used this to inform what coding categories are deployed for not only tweet content, but profiles and other metadata.

CONCLUSION

This chapter makes a case for reflexive, open methods for studying tweets and their users. Importantly, the chapter seeks to emphasize the role of mixed methods for social media research. For example, if social media content and their users are coded by methods of convenience or in biased or unsystematic ways, this has real impacts on the epistemologies presented within the still emergent fields of social media research. This chapter has highlighted the limitations of traditional inductive and deductive methods and underscored some of the potential benefits of alternative approaches such as abductive methods. Retroductive methods and the specific case of grounded theory have been introduced to provide alternative frameworks for studying Twitter data.

Social media are complex sociotechnical spaces. The presentation of the self is often highly nuanced – a case particularly complicated with uses of humor, a frequent theme on Twitter (see Nevin et al., Chapter X, this volume). Coded content can present different perspectives on social interactions, but these data are complementary to computational methods. Combining emergent grounded theory with machine learning or vice versa can complement qualitative and quantitative methods as well as the insights gained from analysis. Such methods can offer new social media ontologies and epistemologies, which pave the way for completely new lines of knowledge.

It is tempting to simply look at easily collectible sets of tweets and make quick observations. However, having methods to systematically and rigorously study tweets produces robust methods as well as new ways to study Twitter data. This chapter has argued that traditional approaches can be useful in studying Twitter, but that alternative approaches, such as retroduction and grounded theory, are of tremendous value to studies of Twitter. Using the #accidentalracist hashtag as a case study, this chapter has presented exemplars of frameworks and methods that I have employed on several Twitter-based projects. These methods range from simple changes to make Twitter research more reflexive and open, to more advanced machine learning approaches. Additionally, employing reflexive ontologies provides ways to see Twitter data from varied perspectives, ultimately advancing our potential to produce more varied and robust social knowledge.

Though computational approaches such as machine learned methods of studying social media content will continue to be important empirical methods, the utility of mixed methods is that they present different perspectives on social interactions within social media. For example, understanding sarcasm within hashtags in tweets is far from straightforward, but content emergently coded by research teams can then be used for supervised learning within traditional machine learning approaches in computer science. The argument here is that mixed methods such as these are fundamentally important to continuing advances in social media research methods, as sometimes very large generalizations are made from Twitter data and this may be a trend.

REFERENCES

Abdul Reda, A., Sinanoglu, S., & Abdalla, M. (2021). Mobilizing the masses: Measuring resource mobilization on Twitter. *Sociological Methods & Research*. doi:10.1177/0049124120986197

Ahn, J., Jeon, Y. A., & Murthy, D. (2020). Differences in health-related social media usage by organizations and individuals. *Telemedicine and e-Health, 26*(6), 812–820. doi:10.1089/tmj.2019.0128

Al-Laith, A., Shahbaz, M., Alaskar, H. F., & Rehmat, A. (2021). AraSenCorpus: A semi-supervised approach for sentiment annotation of a large Arabic text corpus. *Applied Sciences, 11*(5), 2434. https://doi.org/10.3390/app11052434

Alexandre, I., Jai-sung Yoo, J., & Murthy, D. (2021). Make tweets great again: Who are opinion leaders, and what did they tweet about Donald Trump? *Social Science Computer Review*. doi:10.1177/08944393211008859

Baader, F. (2003). *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge: Cambridge University Press.

Balakrishnan, V., Khan, S., & Arabnia, H. R. (2020). Improving cyberbullying detection using Twitter users' psychological features and machine learning. *Computers & Security, 90*, 101710. https://doi.org/10.1016/j.cose.2019.101710

Barker-Plummer, B., & Barker-Plummer, D. (2017). Twitter as a feminist resource: #YesAllWomen, digital platforms, and discursive social change. In J. Earl, & D. A.

Rohlinger (eds), *Social Movements and Media* (Vol. 14, pp. 91–118): Emerald Publishing Limited.

Bhaskar, R. (1976). *A Realist Theory of Science*. Brighton: The Harvester Press.

Blaikie, N. (2004). Retroduction. In M. S. Lewis-Beck, A. Bryman, & T. F. Liao (eds), *The SAGE Encyclopedia of Social Science Research Methods* (pp. 973). Thousand Oaks: Sage.

Bollen, J., Pepe, A., & Mao, H. (2009). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *arXiv preprint arXiv:0911.1583*.

Bornmann, L., Haunschild, R., & Patel, V. M. (2020). Are papers addressing certain diseases perceived where these diseases are prevalent? The proposal to use Twitter data as social-spatial sensors. *PloS one, 15*(11), e0242550. https://doi.org/10.1371/journal.pone.0242550

boyd, d., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15*(5), 662–679. doi:10.1080/1369118x.2012.678878

Bryman, A., & Burgess, B. (2002). *Analyzing Qualitative Data*, London: Routledge.

Charmaz, K., & Mitchell, R. G. (2001). Grounded theory in ethnography. In P. Atkinson, A. Coffey, S. Delamont, J. Lofland, & L. Lofland (eds), *Handbook of Ethnography* (pp. 160–174). London: Sage.

Chen, E., Deb, A., & Ferrara, E. (2021). #Election2020: The first public Twitter dataset on the 2020 US Presidential election. *Journal of Computational Social Science*. doi:10.1007/s42001-021-00117-9

Chen, K., Duan, Z., & Yang, S. (2021). Twitter as research data: Tools, costs, skill sets, and lessons learned. *Politics and the Life Sciences*, 1–17. doi:10.1017/pls.2021.19

Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review, 51*(4), 661–703.

Corbin, J., & Strauss, A. (2015). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Los Angeles: Sage.

Cousineau, L. S., Oakes, H., & Johnson, C. W. (2019). Appnography: Modifying ethnography for app-based culture. In D. C. Parry, C. W. Johnson, & S. Fullagar (eds), *Digital Dilemmas: Transforming Gender Identities and Power Relations in Everyday Life* (pp. 95–117). Cham: Springer International Publishing.

Dann, S. (2010). Twitter content classification. *First Monday 15*(12). http://dx.doi.org/10.5210/fm.v15i12.2745

Devaraj, A., Murthy, D., & Dontula, A. (2020). Machine-learning methods for identifying social media-based requests for urgent help during hurricanes. *International Journal of Disaster Risk Reduction, 51*, 101757. https://doi.org/10.1016/j.ijdrr.2020.101757

Duggan, M. (2017). Questioning 'digital ethnography' in an era of ubiquitous computing. *Geography Compass, 11*(5), e12313. https://doi.org/10.1111/gec3.12313

Ellenberg, H., & Klötzli, F. (1972). *Waldgesellschaften und Waldstandorte der Schweiz*. Institut de recherches forestières.

Essence Magazine. (2013, April 12). Essence poll: Are you offended by 'accidental racist'? *Essence*. Retrieved from http://www.essence.com/2013/04/12/essence-poll-are-you-offended-accidental-racist

Florini, S. (2014). Tweets, tweeps, and signifyin': Communication and cultural performance on 'Black Twitter'. *Television & New Media, 15*(3), 223–237. doi:10.1177/1527476413480247

Glaser, B. G., & Strauss, A. L. (2009). *The Discovery of Grounded Theory: Strategies for Qualitative Research*: Transaction Publishers.

Greenhow, C., Staudt Willet, K. B., & Galvin, S. (2021). Inquiring tweets want to know: #Edchat supports for #RemoteTeaching during COVID-19. *British Journal of Educational Technology, 52*(4), 1434–1454. https://doi.org/10.1111/bjet.13097

Gross, A., & Murthy, D. (2014). Modeling virtual organizations with Latent Dirichlet Allocation: A case for natural language processing. *Neural Networks, 58*, 38–49. https://doi.org/10.1016/j.neunet.2014.05.008

Hardt, M., & Negri, A. (2005). Multitude war and democracy in the age of Empire. Retrieved from www.overdrive.com/search?q=F26C8E99-A346-4C2E-A581-8cda0401d414

Harré, R. (1976). The constructive role of models. In L. Collins (ed.), *The Use of Models in The Social Sciences* (pp. 16–33). London: Tavistock Publications.

Hernandez, L. A. R., Callahan, T. J., & Banda, J. M. (2021). A biomedically oriented automatically annotated Twitter COVID-19 dataset. *ArXiv*. https://arxiv.org/abs/2107.12565

Hernandez-Suarez, A., Sanchez-Perez, G., Toscano-Medina, K., Perez-Meana, H., Portillo-Portillo, J., Sanchez, V., & García Villalba, L. J. (2019). Using Twitter data to monitor natural disaster social dynamics: A Recurrent neural network approach with word embeddings and kernel density estimation. *Sensors, 19*(7), 1746. https://doi.org/10.3390/s19071746

Hogan, B., & Quan-Haase, A. (2010). Persistence and Change in Social Media. *Bulletin of Science, Technology & Society, 30*(5), 309–315. doi:10.1177/0270467610380012

Honeycutt, C., & Herring, S. C. (2009). Beyond microblogging: Conversation and collaboration via Twitter. Paper presented at the *System Sciences, 2009. HICSS'09 42nd Hawaii International Conference*.

Jha, V. K., Hrudya, P., Vinu, P. N., Vijayan, V., & Prabaharan, P. (2020). DHOT-repository and classification of offensive tweets in the Hindi language. *Procedia Computer Science, 171*, 2324–2333. https://doi.org/10.1016/j.procs.2020.04.252

Kumar, A., & Jaiswal, A. (2020). Systematic literature review of sentiment analysis on Twitter using soft computing techniques. *Concurrency and Computation: Practice and Experience, 32*(1), e5107. https://doi.org/10.1002/cpe.5107

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? *Proceedings of the 19th international Conference on World Wide Web* (pp. 591–600). https://doi.org/10.1145/1772690.1772751.

Lindgren, S., & Lundström, R. (2011). Pirate culture and hacktivist mobilization: The cultural and social protocols of #WikiLeaks on Twitter. *New Media & Society, 13*(6), 999–1018. doi:10.1177/1461444811414833

Manovich, L. (2001). *The Language of New Media.* Cambridge: MIT Press.

Merkley, E., Bridgman, A., Loewen, P. J., Owen, T., Ruths, D., & Zhilin, O. (2020). A rare moment of cross-partisan consensus: Elite and public response to the COVID-19 pandemic in Canada. *Canadian Journal of Political Science, 53*(2), 311–318. doi:10.1017/S0008423920000311

Murphy, T. S. (2001). Ontology, deconstruction, and empire. *Rethinking Marxism, 13*(3–4), 16–23.

Murthy, D. (2011a). Emergent digital ethnographic methods for social research. In S. Hesse-Biber (ed.) *Handbook of Emergent Technologies in Social Research* (pp. 158–179). Oxford: Oxford University Press.

Murthy, D. (2011b). Twitter: Microphone for the masses? *Media, Culture & Society, 33*(5), 779–789. doi:10.1177/0163443711404744

Murthy, D. (2018). *Twitter*. Cambridge: Polity Press.

Murthy, D., Gross, A., & Pensavalle, A. (2016). Urban social media demographics: An exploration of Twitter use in major American cities. *Journal of Computer Mediated Communication, 21*(1), 33–49.

Murthy, D., & Gross, A. J. (2017). Social media processes in disasters: Implications of emergent technology use. *Social Science Research, 63*, 356–370. https://doi.org/10.1016/j.ssresearch.2016.09.015

Murthy, D., & Sharma, S. (2019). Visualizing YouTube's comment space: Online hostility as a networked phenomena. *New Media & Society, 21*(1), 191–213. doi:10.1177/1461444818792393

Muse, H. (2013, April 9). 'Accidental Racist' lights up Twitter. *Fortune*. Retrieved from fortune.com/2013/04/09/accidental-racist-lights-up-twitter/

O'Neal, A., Rodgers, B., Segler, J., Murthy, D., Lakuduva, N., Johnson, M., & Stephens, K. (2018). Training an emergency-response image classifier on signal data*. Paper presented at the 2018 17th IEEE International Conference on Machine Learning and Applications, pp. 751-756*.

Olsen, W. K. (2012). *Data Collection: Key Debates and Methods in Social Research*. London, Thousand Oaks, CA: Sage.

Pandya, A., Oussalah, M., Kostakos, P., & Fatima, U. (2020). *MaTED: Metadata-assisted Twitter Event Detection system*, Cham: Springer.

Paul, G. (1993). Approaches to abductive reasoning: An overview. *Artificial Intelligence Review, 7*(2), 109–152. doi:10.1007/bf00849080

Petersen, K., & Gerken, J. M. (2021). #Covid-19: An exploratory investigation of hashtag usage on Twitter. *Health Policy, 125*(4), 541–547. https://doi.org/10.1016/j.healthpol.2021.01.001

Poole, S. (2015). Can computational approaches motivate new theories? *Computational Approaches to Advance Communication Research Pre-Conference, International Communications Association Annual Conference*.

Quan-Haase, A., Martin, K., & McCay-Peet, L. (2015). Networks of digital humanities scholars: The informational and social uses and gratifications of Twitter. *Big Data & Society, 2*(1). doi:10.1177/2053951715589417

Rafail, P. (2018). Nonprobability sampling and Twitter: Strategies for semibounded and bounded populations. *Social Science Computer Review, 36*(2), 195–211. doi:10.1177/0894439317709431

Robertson, B. W., Johnson, M., Murthy, D., Smith, W. R., & Stephens, K. K. (2019). Using a combination of human insights and 'deep learning' for real-time disaster communication. *Progress in Disaster Science, 2*, 100030. https://doi.org/10.1016/j.pdisas.2019.100030

Samper-Escalante, L. D., Loyola-González, O., Monroy, R., & Medina-Pérez, M. A. (2021). Bot datasets on Twitter: Analysis and challenges. Applied Sciences, 11(9), 4105. https://doi.org/10.3390/app11094105

Sloan, L., Jessop, C., Al Baghal, T., & Williams, M. (2020). Linking survey and Twitter data: Informed consent, disclosure, security, and archiving. *Journal of Empirical*

Research on Human Research Ethics, 15(1–2), 63–76. doi:10.1177/1556264619853447

Stock, W. A., Benito, J. G., & Lasa, N. B. (1996). Research synthesis: Coding and conjectures. *Evaluation & the Health Professions, 19*(1), 104–117. doi:10.1177/016327879601900108

Webb, H., Jirotka, M., Stahl, B. C., Housley, W., Edwards, A., Williams, M., Procter, R, Rana, O., & Burnap, P. (2017). The ethical challenges of publishing Twitter data for research dissemination. *Websci'17: Proceedings of the 2017 ACM on Web Science Conference* (pp. 339–348). https://doi.org/10.1145/3091478.3091489

Wu, S., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Who says what to whom on Twitter. *Proceedings of the 20th International Conference on World Wide Web, pp 705-714.* http://dx.doi.org/10.1145/1963405.1963504

Xin, E. Z., Murthy, D., Lakuduva, N. S., & Stephens, K. K. (2019). Assessing the stability of tweet corpora for hurricane events over time: A mixed methods approach. *Proceedings of the 10th International Conference on Social Media and Society, pp 59-66*. https://doi.org/10.1145/3328529.3328545

Zhang, C., Xu, S., Li, Z., & Hu, S. (2021). Understanding concerns, sentiments, and disparities among population groups during the COVID-19 pandemic via Twitter data mining: Large-scale cross-sectional study. *Journal of Medical Internet Research, 23*(3), e26482. doi:10.2196/26482

Zimmer, M., & Proferes, N. J. (2014). A topology of Twitter research: Disciplines, methods, and ethics. *Aslib Journal of Information Management, 66*(3), 250–261. doi:doi:10.1108/AJIM-09-2013-0083