This post will be on the burgeoning social media site Twitter. One of the main differences between old and new methods is that Twitter is both a medium of research, but also a recruitment tool. Twitter is a micro-blogging site that allows people to send a message of a 140 characters or less. These messages are called tweets and are generally public although your account can be set to private (Hughes & Palen 2009). It is a platform that allows people to organically voice and share their own personal lives. Twitter as a social media site allows researchers access to their entire archived back catalogue of publicly available tweets over the last nine years (Tornes 2015). This is achieved via an API (Application Programming Interface) that is
essentially a structured way of having access to a particular piece of data. On average there are 500 million tweets sent per day (Twitter 2015). This is an almost unprecedented amount of longitudinal data to highlight emerging trends and popular thought. You are also able to generate data through the analysis of differences in tweets across countries and cultures (Burgess & Brun 2012). You access this plethora of data via applications that have access to Twitter’s systems. Applications such as DiscoverText and Nuvi are specialists in granting access to this data for researchers (DiscoverText 2015; Nuvi 2015).

This method of data collection is very powerful and exactly how it can be applied is still being investigated. There are a diverse set of research projects that have utilised Twitter has a platform to collect data. One such study is Hughes and Palen who utilised Twitter to collect data on emergency events (2009). Another utilised Twitter to predict flu trends to help reduce the impact of seasonal flu epidemics (Achrekar, Lazarus & Park 2011). From a personal perspective I am interested in the change of social norms through popular culture. Twitter would allow me to utilise publicly available tweets to generate data on the trends of social norms across a timespan. More specifically the trend in alternative sexual lifestyles and the impact that popular culture has had on inclusiveness and acceptance. Twitter, as opposed to more traditional methods would allow me to relatively inexpensively collect a large amount of data to determine whether any changes have occurred.

There are of course many different strengths and weaknesses to any research method, and using Twitter as a data collection tool is no different. The major strength of using Twitter and other social media sites is the amount of data available to researchers. For example Twitter alone is reported to have 319 million active monthly users (Twitter 2015). This is a phenomenal amount of data that is able to be leveraged by researchers to investigate social phenomena. The sheer amount of data available can also be a weakness as researchers are overwhelmed by the deluge of data (Burgess & Bruns 2012). This strength is particularly highlighted by how accessible Twitter data is. Tweets are generally of a uniform length and therefore
easily transformable into readable data. This is paired with the prevalent use of hashtags that allow geographically distant people to have one conversation about a singular event. These are some of the reasons why Twitter is such a popular platform for researchers in comparison to social media sites such as Facebook, and Instagram. This moves into a general weakness of Twitter as well as all social media sites, that of access. Access to data is generally filtered, and in the case of Twitter heavily regulated (Burgess & Bruns 2012). Who has access to this data, and what is done with it before researchers get it? These are some of the questions that researchers need to ask before embarking on using social media data. Twitter for example has two available avenues of access to its archive of tweets. One is its firehose API, which has the name suggests gives access to the unfiltered pipeline of public tweets (Burgess & Bruns 2012; Tornes 2015). The other is that of a streaming API, which means that only a ‘small’ percentage of the overall tweets are utilised. This can have issues for researchers to the integrity of the data they are using.

In conclusion, while this has been a shallow overview of using Twitter as a data collection tool, it is obvious that there are many positives as well as a few negatives. Though in my opinion this access to near real-time data is invaluable and when treated with caution is an extremely powerful tool.

References


Recently Added Posts:

- **Resilience in Gaming Survey**
  - August 10, 2016
  - In "Data Science"

- **Open Data Repositories**
  - March 22, 2016

- **Microsoft OneNote**
  - March 18, 2016
Expert panel: the state of the National Broadband Network
(http://www.transcendingdigital.space/expert-panel-the-state-of-the-national-broadband-network/)

MARCH 14, 2016

Importing data into R (http://www.transcendingdigital.space/importing-data-into-r/)

FEBRUARY 29, 2016