Cassie Chong Kenci

Varshini Harthachitramalogan

*Introduction to Digital Humanities, Jan-Apr 2022 term*

## Background

Content consumption before the Internet was done mainly from traditional media such as television, radio, and print media. Hence, content generation fell largely in the hands of media companies. However, with the rise of the personal computer, the Internet, and personal mobile devices, we now have a digital world that is globalised and interconnected.

This has several consequences on how content is handled. Firstly, the consumption of media is now not limited to the use of traditional media and word-of-mouth. Secondly, the production of content is not constrained to media corporations or writers; every user of the Internet has the ability and means to create and spread content for the whole digital world to scrutinise.

With this increase in accessibility to creating and sharing content, it is a given that the creator economy has taken off. Over 50 million people claim to be content creators, sharing and creating content onto the Internet, achieving goals such as monetisation or influential status (Yuan & Constine, 2021). A greater number of users are also consuming user-generated content at a greater proportion as compared to traditional media (Lentini, 2022).

Because content on the Internet has a potentially great influence on the opinions of its consumers, it is imperative to know what kind of content holds such influential power. It also becomes of importance to explore how the nature of content shared and consumed relates to the consequences it may bring to its consumers.

## Purpose

The purpose of this proposal is to learn the role of digitally spread content in eliciting emotional reactions from its consumers. Additionally, the proposal seeks to discover the impacts such content can have on the opinions of its consumers. We hope that through a greater understanding of the impact of words and content unto others, that people use words, create, and share content while being mindful of its potential impact on those who read and receive them.

## Objectives

The objective of this proposal is to identify the various topics that stir the greatest number of responses from the public, namely in shares and comments. Shares implies that the consumer has decided to spread this content further to their social networks, deciding that this is a worthy piece of content that either aligns or disagrees strongly with their views, or has content that led them to decide this should be spread to their social circles in some way or another. Comments will tell the nature of the consumer's reaction, be it emotional, approval or disapproval, comical or disturbing.

By analysing the topics which stir the greatest responses, we can conclude the nature of content that the public reacts the most strongly to. Further analysis on the nature of such reactions will tell if that content elicits a positive or a negative response. Doing so will allow us to deduce the macro impact of digital content on its consumers.

## Significance

The accessibility of content and content-sharing methods facilitate online activism, particularly by the documentation of individual opinions and the sharing of common stands (Greijdanus et al., 2020). Activities such as online petitions, crowdsourcing, trendsetting and more can gain traction in the digital world, serving as a call to action for its consumers in achieving their respective goals, be it for profit or for influence.

It is important that we acknowledge the potential impact this can bring, especially when such content leads to undesired consequences, regardless of whether it is intentional or not. For instance, the recent events between Russia and Ukraine have led to an outpour of support for Ukraine, but also stirred antagonism towards Russians. This has unintended effects on even innocent Russians far away from home, such as the owners of a Russian dumpling restaurant in Singapore receiving hate comments (Abdullah, 2022). The intentional use of digital content to spread hate is also plausible, exemplified by terrorists using social media platforms for radicalisation, recruitment, funding, planning and execution of terror activities (Interpol, 2019). Closer to home, several lone attack and terror plots have been fortunately discovered and foiled by the authorities in Singapore in recent years. (MHA, 2021).

Hence, there is a need for us to understand, in this world where content is easily created, accessed, and spread, how content holds the power to influence people.

## Review of *Content-Era Ethics (McNulty, Tess. 2021)*

This proposal was inspired by the work of McNulty, an academic who, in this article, delved into the ethical examinations of viral content. She aims to examine how social media is reshaping popular ethics. Her thesis states that social media content, with its pursuit of "prosumption", promotes heightened moral concern, but simultaneously reduces moral action to aesthetic self-expression.

First, she identifies features of content that encourage “prosumption”, which she defined as “consumption” as expressive “production”. With topic modelling applied to an archive of viral content, McNulty identified three categories which constitute viral content, namely real sensational events, the self, and artifice or crafts. Then, she used word frequency analyses and the NRC Lexicon to associate emotion with the topics, and concluded that the first category was the most prominent. From which, it was also found that sentimental media, or what McNulty calls the "uplifting anecdote", was the genre of highest prevalence in viral content.

Moving on, McNulty used a trained Naïve Bayes classifier on the headlines, which confirmed that sentimental media invoked the ethical in one of two ways: inspiring sympathy or depicting heroism, or by compelling divisive responses. The tool observed the word frequency in each type of headline, suggesting that the first way was the most prominent. The tool also provided many insights which McNulty used to conclude that the genre "uplifting anecdote" consistently suggests that all moral action occurs on the plane of symbolic self-expression, and hence implies that the predominant ethical crime is not inaction, but false self-depiction. She then argues that most ethical debate about content revolves around the authenticity of the players in the content.

Finally, she briefly examined two works that are "uplifting anecdotes", which are also allegories for social media. These two works are NBC's hit sitcom The Good Place and George Saunder's novel Lincoln in the Bardo. She concludes her article with her statement that the social media era has created a world in which "all actions are symbolic and expressive, and the measure of its quality is its perceived authenticity". There are two types of stances one can take with this conclusion. One insists that ethical action is thus pointless, while the other strives to prove otherwise.

No doubt, McNulty's approach to her thesis is well-supported by her use of text tools and critical analyses. However, it appears that her usage of text tools has narrowed down her data pool. From her original aim of showing how social media is reshaping popular ethics, she ended up narrowing down to social media content of the uplifting and sentimental nature, citing reasons that these are statistically more popular and thus play a greater role in "reshaping popular ethics". We feel that by narrowing down on one genre of popular content, she has removed some potential insights into analysing the other categories in greater detail.

Additionally, while roping ethics into play, McNulty's analyses were limited to the moral scope of the impact of content on its consumers, and the moral constituents of the content itself. She focused on how the incorporation of morality into content fuelled its prosumption, and also used an ethical standpoint to evaluate the consequences of the pursuit of prosumption in social media. While this gave an in-depth understanding of the ethical implications regarding the proliferation of morally associated content on social media and its impact on its consumers, this also prompted us to think of the impact that social media content can bring outside the ethical realm.

Hence, we wish to approach our proposal unconstrained by category or the ethical domain and explore how various types of content inspire reactions and shape opinions of the audience, which could have an impact on the real world in ways not just ethical, but in other aspects such as economic, political, et cetera. Even though the scope of analysis has changed, the analysis methods employed by McNulty could still serve as a great example for our approach.

## Research Questions of Proposed Project

The questions this project aims to tackle are: What are the types of text content that users engage with most? By engagement, we refer to social media content that has a high number of comments and shares. What type of content receives more positive or negative comments? How does the content influence and impact real-time opinions of the audience and what kind of potential ramifications does this have on society?

## Proposed Methodology

We require two main findings: what constitutes influential content, and what influence they bring. Therefore, for this proposal, we split the analysis into two portions to suit our needs.

Referring to McNulty's method of first using topic modelling, then word frequency analysis, then sentiment analysis, we can also use a similar approach.

Although we wish to have a general view of all content, it is labour-intensive to look at all social media platforms. We assume that various platforms have similar content and reactions, and choose one platform to focus on: Twitter, for its ease of "retweeting" and holding reply threads to posts. The nature of its posts, being microblogs, also encourage users to express their opinions in a concise manner and allows consumers of the content to access information in great quantities and offer their near-immediate responses. Due to the nature of Twitter being an unstructured online social network, we need both quantitative and qualitative analyses on its data to obtain information. Quantitative analyses will include numerical metrics such as the number of "retweets" and replies which will point to how the thread of content is spread. Qualitative analyses will include topic modelling and sentiment analysis on what content is made up of, and the nature of the responses to content.

It is more direct to take a backward approach, in which we first look at the reactions of the users before exploring what caused them. This is because replies and comments take up a far greater number than its original posts. First, we extract Twitter posts from the last two years using the Twitter public API (Twitter, 2022) and store them in an accessible database. The API is able to allow us to filter Tweets by its attributes, accounts of interest, contents, and Tweet type. It can also filter away Tweets that are meaningless in our research such as advertisements and Tweets by bots (Twitter, 2022). It is here we can separate the "retweet" and "reply" types from Tweets coming from users and "verified" accounts. The "retweet" and "reply" types are taken to be in reaction to Tweets coming from users and "verified" accounts, so we have two categories, the reactions and the content created.

We perform a sentiment analysis to the reaction category. First, we perform a word frequency analysis using text tools, such as Voyant (Sinclair & Rockwell, 2016). This will identify the most frequently used terms, whether positive or negative, among reactions to content. We further use the NRC Lexicon employed also by McNulty to separate these words into eight basic emotions. This will allow us to classify the kind of reaction these Tweets contain towards the content they were reacting to. Additionally, we can also use the DESSERT (Dahiya et al., 2021), a novel deep state-space model that leverages the function approximation capability of deep neural networks to classify hate speech, and will process all the replies in the Tweet-reply chains into positive or negative replies.

Now that we have our reactions classified into what type of responses they are, we can attempt to link the responses back to what had caused them. For each category of response, we perform topic modelling, with a classification tool such as MALLET (McCallum, Andrew Kachites, 2002). We can then discover the various topics which spurred that particular type of response. Doing this across different categories of responses, we may also find certain topics which cause multiple, or even conflicting responses among different users.

The topics which elicit singular kinds of responses can be identified to be those which the masses generally react similarly to. Together with the kind of response they elicit, conclusions can be made about the topic's nature and what kind of response the masses give such content. The topics with multiple kinds of responses which may conflict may point to controversial topics which spur debate. It is also possible to analyse, from the quantity of responses, which kind of response is in the majority, and which are not. Conclusions can then be made about which kind of content is controversial and spur debate rather than unified support or disapproval.

## Research Limitations

As mentioned earlier, several assumptions are made in our approach. To assume that various social media platforms have similar content and reactions may be a pitfall, as different platforms, with their different forms of communication, may lead to different content and responses being published. For example, Facebook's focus on personal networking may lead to responses which are more curated to the users' social circles, in comparison to Twitter, which focuses more on sharing succinct snippets of information per topic regardless of personal social connections. This may have removed a significant portion of the type of responses and content which are spread on the Internet.

With the focus on Twitter, we are also narrowing down on the demographics of the users of Twitter and excluding those who prefer other platforms. For instance, as of 2020, the majority of users on Twitter are in the 25-34 years old age category and are male (Sehl, 2020). Other demographic indicators, such as access to the Internet and the popularity of Twitter as a social media platform in each country, may lead to a skew in the data towards the backgrounds of users which use Twitter more. Thus, the results that we get may be biased by these demographic data and are not able to represent all users of social media across all platforms fairly.

In addition, due to our approach using text tools, we are also neglecting another communication form of the digital era: pictures and videos. With the rise of Instagram and TikTok, we cannot ignore the significance of content being created and shared across via these mediums. However, to include multimedia content into our dataset, we will need additional tools for image and speech recognition to ascertain the content of the media before analysing them accordingly. There is also the risk that such media can be interpreted differently, or even misinterpreted among different users, which increases the difficulty of using tools to determine the contents of images and videos.

## Conclusion

In conclusion, we hope that this project can establish some degree of understanding of what content can influence its consumers, and in what ways. Being able to draw conclusions about the impact of content on its consumers, various stakeholders can be more informed in how content is used and how it can influence others. Content creators can be more cognisant of how their content can impact their consumers. This can help them curate or design their content to better suit their needs. On the other hand, consumers of content can be more aware of how certain topics can create unified or divisive stances, or how content can make them feel a certain way, and hence be more mindful of the kind of content they consume and decide to comment and share. It is hoped that such mindfulness can create a better online community for all to share.

## Bibliography

Yuan, Y., & Constine, J. (2021, November 9). *What is the creator economy? influencer tools and Trends*. SignalFire. Retrieved March 5, 2022, from <https://signalfire.com/blog/creator-economy/>

Lentini, N. (2022, January 6). *User-created content - especially video - gains on traditional media consumption*. Marketing Dive. Retrieved March 5, 2022, from <https://www.marketingdive.com/news/user-created-content-especially-video-gains-on-traditional-media-consum/616745/>

*Analysing social media*. INTERPOL. (n.d.). Retrieved March 5, 2022, from <https://www.interpol.int/Crimes/Terrorism/Analysing-social-media>

*Singapore Terrorism Threat Assessment Report 2021*. Ministry of Home Affairs. (2021, June 23). Retrieved March 5, 2022, from <https://www.mha.gov.sg/mediaroom/press-releases/singapore-terrorism-threat-assessment-report-2021/>

Abdullah, S. (2022, March 2). *No 'hate speech': Russian dumpling restaurant in Singapore asks people to be 'kinder in their words'*. CNA Lifestyle. Retrieved March 5, 2022, from <https://cnalifestyle.channelnewsasia.com/dining/russian-dumpling-restaurant-singapore-dumplings-ru-ukraine-305971>

Greijdanus, H., Fernandes, C. A. de M., Turner-Zwinkels, F., Honari, A., Roos, C. A., Rosenbusch, H., & Postmes, T. (2020, March 21). *The psychology of online activism and social movements: Relations between online and Offline Collective Action*. Current Opinion in Psychology. Retrieved March 5, 2022, from <https://www.sciencedirect.com/science/article/pii/S2352250X20300324>

McNulty, Tess. 2021. “Content-Era Ethics.” *Journal of Cultural Analytics* 6 (2). <https://doi.org/10.22148/001c.22220>.

Twitter. (n.d.). *Getting historical tweets using the full-archive search endpoint | docs | twitter developer platform*. Twitter. Retrieved March 10, 2022, from <https://developer.twitter.com/en/docs/tutorials/getting-historical-tweets-using-the-full-archive-search-endpoint>

Twitter. (n.d.). *Learning path: How to detect signal from noise and build powerful filtering rules | docs | twitter developer platform*. Twitter. Retrieved March 10, 2022, from <https://developer.twitter.com/en/docs/tutorials/building-powerful-enterprise-filters>

Sinclair, S., & Rockwell, G. (2016). *Voyant tools*. Voyant Tools. Retrieved March 10, 2022, from <https://voyant-tools.org/>

Mohammad, S. (2011). *NRC Word-Emotion Association Lexicon*. NRC Emotion Lexicon. Retrieved March 10, 2022, from <https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm#:~:text=The%20NRC%20Emotion%20Lexicon%20is,were%20manually%20done%20by%20crowdsourcing.>

Dahiya, S., Sharma, S., Sahnan, D., Goel, V., Chouzenoux , E., Elvira, V., Majumdar, A., Bandhakavi, A., & Chakraborty, T. (2021, August 1). *Would your tweet invoke hate on the fly? forecasting hate intensity of reply threads on Twitter: Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining*. ACM Conferences. Retrieved March 10, 2022, from <https://dl.acm.org/doi/pdf/10.1145/3447548.3467150>

McCallum, Andrew Kachites. (2002) *"MALLET: A Machine Learning for Language Toolkit."* <http://mallet.cs.umass.edu>

Sehl, K. (2020, July 14). *Top twitter demographics that matter to social media marketers*. Social Media Marketing & Management Dashboard. Retrieved March 10, 2022, from <https://blog.hootsuite.com/twitter-demographics/>