Collaboration of Open Content News in Wikipedia: The Role and Impact of Gatekeepers

Ang Li
University of Pittsburgh
Pittsburgh, PA
ANL125@pitt.edu

Rosta Farzan
University of Pittsburgh
Pittsburgh, PA
rfarzan@pitt.edu

ABSTRACT
With the advances of social technologies, open content in social media has become an important place where people gather and communicate news. Prior studies in conventional mass media has long studied the news reporting process, and suggest that gatekeepers – editors or journalists who control the information and have a unique power to determine what gets published to the public – play an important role in the news reporting process. However, as the process of how open content are created by contributors in social media platforms is different, what we understand about content publication process in traditional mass media can not directly apply in the context of social media. Especially, it is unclear who are the gatekeepers and how do they influence the content creation and spread of information in social media. In the current proposed study, I aim to understand this new model of content generation process through the lens of gatekeepers in social media platforms such as Wikipedia. Specifically, I aim to discover ways to identify gatekeepers and assess their impact on information quality and content polarization.

CCS CONCEPTS
• Information systems → Users and interactive retrieval.

ACM Reference Format:

1 INTRODUCTION
A series of major protests are happening in Hong-Kong. Hurricane Dorian causes 20 deaths in the Bahamas. The United Kingdom is withdrawing from the European Union. Events unfold as our days go on and attract attentions of many individuals across the world. In facing such breaking news events, people have traditionally relied on mass media to seek updated information about breaking news events by turning on the TV, listening to the radio, or checking the newspaper. Nowadays with the advantage of social technologies, there has been a significant shift in ways that people gather and communicate news, and increasingly open content in social media has become an important source of news. Based on a 2018 survey conducted by the PEW research center, 25% of the U.S. adults get their news via social media – this number is even slightly higher than the share who often do so from print newspapers (16%).1

Rich prior literature[19, 23] has long studied the news reporting process in conventional mass media (i.e. print, radio, TV), and provided an understanding of the journalistic news gathering and news publishing process in the mass media. As the conventional mass media has their inherent structure in the availability of column space, air time, or transmission frequencies[3], it is impossible for media outlets to present all information to audiences. Certain mechanisms or processes are established so that media outlets can control and select the information to be published according to specific criteria of newsworthiness[8]. However, the editors or journalists, known as media gatekeepers, may suppress or skew information in the process as news outlets usually have their own ideological preferences[5, 18, 23]. As a result, the skewed information can influence the perception of citizens about the relevant and importance of the news events[5, 30]. The Gatekeeping theory thus studies and investigates this mechanisms or processes through which the media control the construction and dissemination of information.

Gatekeeping is defined as the process by which "the billions of messages that are available in the world get cut down and transformed into the hundreds of messages that reach a given person on a given day[20, 24, 32]." In this process, Gatekeepers – the actors who craft and conduct what is being published to the masses – are critical, as they decide what information the general public gets to view[24, 25, 32]. Gatekeepers not only control the information quality that they deem worthy to be distributed to people, but they also have a unique power to influence audiences by selecting and filtering certain information. As a result, they can determine what is to become the public’s view of the world events with little involvement from the general public’s opinions[24, 27].

In contrast to this traditional process, it has been argued that the process of how open content is created by contributors in social media, and especially in Wikipedia is different. The difference can be contributed to the following three dimensions of content generation in social media. First, unlike traditional mass media, open content in Wikipedia can be created or edited by almost anyone. As a result, the public is not just the content consumers but also the content creators. They have the opportunity to get involved in the content generation process. Second, Wikipedia encourages collaboration among contributors[7], which can facilitate and support a collective determination of what content is created and disseminated. Although the editors are distributed across different

In this study, I propose to focus on the gatekeepers in Wikipedia and to investigate what is their role and how do they impact the process of collaborative content creating?, and particularly, how do the role of gatekeepers and the collaborative open content creation process relate to content polarization?

2 BACKGROUND LITERATURE ON WIKIPEDIA

It has been demonstrated [6, 14, 17, 22, 31] that, when current events happen in the society, peer production platforms like Wikipedia serve as a place for many citizens to seek and share the latest information, to collectively make sense of what is happening, and build memories of the event. Pentzold claims that Wikipedia is a place for a global memory of events and describes its essential role in this dynamic process of constructing collective memories[22]. Especially, the popularity and openness of Wikipedia together with its detailed records enable general population to present and organize different sources, information and diverse perspectives into coherent narratives. A number of case studies have been conducted to understand how editors collaboratively produce knowledge as well as manage conflicts. For example, researchers have investigated the memory-building processes for the Egyptian revolution across languages[6], the deaths of notable people[14], the Vietnam War[17], and the Black Lives Matter social movements[31].

In addition, Wikipedia has long been considered as an open system that facilitates the mass collaboration among contributors. In addition, Wikipedia has strong community principles that all contributors should follow2. Although the Wikipedia editors are distributed across different locations, those who interested in one topic come together under Wikipedia’s principle to improve the same article collaboratively.

3 RESEARCH METHOD

In the current study, I plan to focus on discovering ways to identify gatekeepers in Wikipedia and to quantitatively measure their power of gatekeeping. As discussed before, gatekeepers in the traditional mass media are the newsrooms or journalists who have the absolute power to determine the news content that is presented to the readers (i.e. general public). The traditional news audience is the “gated”, bounded by the scope of events and perspectives presented in coverage – they rely mainly on the gatekeepers to get information and thus their perspectives are also largely influenced by them. However, in the context of online social media where newsrooms or journalists do not hold absolute power of gatekeeping any more, the general news audience (i.e. those being “gated” in traditional mass media) can also be involved in this process and create measurable impact to influence other’s perspectives – they can produce and broadcast their own opinions, to choose various content, and to even interact with traditional gatekeeping newsrooms or journalists and influence them.

The theory of Networked gatekeeping emphasizes this shifted power dynamics in the digital age and describes the relationships between news actors who hold diverse levels of power and positions. Particularly, the theory suggests that the power of “gated” (i.e. 2https://en.wikipedia.org/wiki/Wikipedia:Expectations_and_norms_of_the_Wikipedia_community
news audiences) in the networked gatekeeping process can be described by four attributes – their opinion leadership, information production ability, relationship with traditional gatekeepers, and their autonomy to choose diverse information[2, 33]. Therefore, in the current study, I plan to operationalize the four attributes proposed by Network gatekeeping theory and quantify the power of social media gatekeepers for the open content contributors.

To do so, a mixed-method study including machine learning techniques and qualitative content analysis will be conducted, aiming to not only discover ways in identifying gatekeepers in different social media platforms but also to evaluate the proposed method. Specifically, I will first collect a data corpus that represents news events related open content and their associated open content contributors in Wikipedia. I will then operationalize the four attributes discussed previously (i.e. the opinion leadership, information production ability, relationship with traditional gatekeepers, and their autonomy to choose diverse information) as a set of variables to infer each contributors’ gatekeeping power in Wikipedia.

Before modeling contributors’ power of gatekeeping, one challenge remains: we don’t have a ground truth for each contributor’s gatekeeping power in social media. Following the definition of social media gatekeepers – they are the users who are more powerful in determining the open content that the general audience eventually read in social media, and influencing readers’ perspective. As a result, the content created by powerful gatekeepers in social media should reach and influence more readers, and should stay in the open content for a longer time. Therefore, the influence of the content created by each contributors – how many readers the content can be eventually reach and like, how long time the content still stays and being transmitted – can be a proxy for the outcome of the gatekeeping. I plan to use the content influence as outcome variable of gatekeeping, and check their relationship with the measures of gatekeeping proposed before as independent variable. The hypothesis is that content contributed by users with more gatekeeping powers should reach more readers, and be able to stay in open content for good amount of time. I will test this hypothesis to validate the proposed measures.

To further evaluate the validity of the proposed variables, I plan to also adopt qualitative methods. Especially, ground truth of who are gatekeepers can be obtained through human annotation of a small random sample of open content contributors. In addition, qualitative content analyses can be conducted to further examining and characterizing the identified gatekeepers and their gatekeeping strategies.

4 EXPECTED RESULTS AND DISCUSSION
The current study is expected to identify a group of gatekeepers in Wikipedia who have more influential power to determine the news open content that most audiences read. Future studies can be done to investigate the role of gatekeepers on the process that contributors collaboratively create open content in Wikipedia, especially, to study how does the power of gatekeepers in relation with the open content quality and polarization in document news events.

News events are unique as they are highly popular and usually involve controversial topics. As identified by many prior studies, reporting them can be polarized in traditional mass media as different media outlets hold different preferences or political stands[10, 11, 29], and this polarization has been a source of many debates in recent years. Meanwhile, news events can also attract the diverse interests from the general public across the world to collectively build open content around them in social media platforms. Such diversity can play a critical role in achieving the goal of “Neutral point of view” – one of the Wikipedia’s major principles. The results of current study can help future studies to provide a set of empirical evidence on whether or how the collective model of news related open content creation process can serve as a step towards a solution to this critical issue.

The world has become more divided than ever before, especially as we face controversial events that are so common in our daily lives. A biased perspective can easily be framed by selecting and curating the information that supports that perspective, and the general public is constantly being influenced by the biased information they are exposed to. Recently, many researchers have also noticed this critical issue, and most of their work focus on discovering ways to identify or detect content bias or polarized opinion (i.e. [4, 9, 12, 16]). However, instead of focusing on detecting content bias, my proposed work focuses on understanding the content generation process and the underlying mechanism, which can be more constructive in both understanding the polarization and a solution to it. Without a better understanding of how the content is selected, curated and spread, the society is in danger of getting more polarized to such an extent that reconciliation between opposing perspectives will become even more difficult.

Therefore, I pursue my study as a step towards solving this critical issue by providing guidelines for practitioners and designers of social media systems that aim to present their users a diverse perspective with unbiased information. Overall, the study will have implications for information scientists, technology developers and community practitioners seeking to build social media platforms that can generate inclusive information that is both high quality and represents diverse perspectives.

REFERENCES


