

depths of wikipedia: Understanding cross-platform online attention, content creation, and success

Keywords: Collective attention, Wikipedia, Online navigation, Cross-platform, Novelty

Extended Abstract

“depths of wikipedia” is a popular social media account dedicated to sharing unusual or interesting content from Wikipedia, with wide online following across Twitter, Instagram, and TikTok (760K, 1.2M, 163K followers respectively). In highlighting Wikipedia articles elsewhere on social media, depths of wikipedia can act as a significant driver of traffic to otherwise unexplored areas of the online encyclopaedia. It is thus interesting to consider how a single account, mediated by the algorithmic affordances of a large social media platform, may shape collective attention and influence content elsewhere on the Web. The nature of depths of wikipedia, in posting unusual articles unrelated to current events, makes it an appealing case study for studying cross-platform collective attention without the interference of external factors.

Frequently, studies of collective attention online are single platform, and/or do not consider how users navigate between different websites, or the effects of this interaction. Related works have looked at why general audiences read Wikipedia [5], modelled peaks of attention with associated editing activity in response to unexpected and planned events [1, 2], examined of different social platforms (including Wikipedia) in their combined response to exogenous events [4], or explored wider patterns of cross-platform linking [3]. However, work on explicit cross-platform navigation and its effects is sparing.

The aims of this study are threefold. Firstly, to understand how online content may be considered unusual and how the depths of wikipedia account is able to identify novel content. Secondly, to understand the effects of substantial inflows of traffic to typically unpopular Wikipedia pages, in terms of both page views dynamics and edits to Wikipedia content. Finally, to understand how unusual content on Wikipedia may be popular elsewhere on the Internet by the response (likes/retweets/views) on Twitter.

To study the dynamics and effects of cross-platform collective attention, we collect 18 months of tweets from the depths of wikipedia account (example in figure 1), taking tweet text and image content (transcribing images where appropriate), and data on likes, retweets, and views. The depths of wikipedia Twitter account is selected over the other social media accounts since direct links to the appropriate Wikipedia article can be and are included in each post. We also use information from Wikipedia on article content, page views, hyperlink network structure, edits, and editors for the linked articles that are featured. After processing, the primary dataset of interest consists of 1027 threads, each with an associated tweet image and unique Wikipedia article. In our work, we compare these tweets and articles against all public tweets with a linked Wikipedia article, as well as those tweeted by the official Wikipedia Twitter account from the same time period. The study is also informed by interview responses from the owner of the depths of wikipedia account. We address this research methodologically with text, network, time series, and regression analysis approaches.

Initially, we attempt to quantify article “depth”, or how unusual/unexpected an article is, in several ways; according to normalised term frequency scores relative to the rest of Wikipedia, by volume of page views, and by article PageRank centrality in the hyperlink network. We find

that, 1) Wikipedia articles featured by depths of wikipedia are, as measured by these factors more unusual than other Wikipedia articles posted on Twitter. However, with the exception of the tweeted image text, the featured content is no more unusual than a typical Wikipedia article. 2) An article being featured by depths of wikipedia typically experiences an instantaneous increase of 345 hourly page views (or around 55x), and an average increase in daily edit rate by a factor of 7.8x, as indicated in figure 2. This is greater than that observed for articles featured by the official Wikipedia Twitter account. Moreover, depths of wikipedia has attracted an estimated 443 new account-holding editors who have gone on to make a further 7416 edits across Wikipedia. 3) Under a simple regression model, features of a Wikipedia article are weakly associated with that article proving popular when posted to Twitter.

Overall, the consequences of these findings suggest that social media accounts like depths of wikipedia can have a significant impact on the visibility and development of online content. The depths of wikipedia account successfully identifies unusual articles for social media, though there are evidently many more candidates available in the online encyclopaedia. Alternative avenues, such as meme accounts, can be significant drivers of attention, in the form of page views and edits / (new) editors, to Wikipedia. They can also be particularly effective in targeting less popular pages as compared to typical sources such as official channels, current events, or popular entertainment. Though one must also be conscious of the risks of attracting this attention, and perhaps low quality or bad faith edits that could follow. Finally, whilst initial results on modelling article popularity on social media are mixed, being able to predict this popularity of content from Wikipedia may help more people discover relevant informative articles, and can assist in finding new editing communities for neglected articles.

Results from this work contribute to understanding how Internet users navigate between popular platforms, as well as the effects of sudden targeted influxes of users to particular sites. The work also offers insights for online content creation more widely in computationally identifying interesting content and optimising clickthrough rates. Furthermore, there are clear practical applications in helping to understand how Wikipedia attracts users and edits for the future promotion and development of the site.

References

- [1] Brian Keegan, Darren Gergle, and Noshir Contractor. Hot off the wiki: Structures and dynamics of Wikipedia’s coverage of breaking news events. *American Behavioral Scientist*, 57(5):595–622, 2013.
- [2] Ryota Kobayashi, Patrick Gildersleve, Takeaki Uno, and Renaud Lambiotte. Modeling collective anticipation and response on Wikipedia. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 15, pages 315–326, 2021.
- [3] Florian Meier. TWikiL—the Twitter Wikipedia link dataset. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, pages 1292–1301, 2022.
- [4] Miles Osborne, Saša Petrovic, Richard McCreadie, Craig Macdonald, and Iadh Ounis. Bieber no more: First story detection using Twitter and Wikipedia. In *SIGIR 2012 Workshop on Time-aware Information Access*, pages 16–76. Citeseer, 2012.
- [5] Philipp Singer, Florian Lemmerich, Robert West, Leila Zia, Ellery Wulczyn, Markus Strohmaier, and Jure Leskovec. Why we read Wikipedia. In *Proceedings of the 26th International Conference on World Wide Web*, pages 1591–1600, 2017.

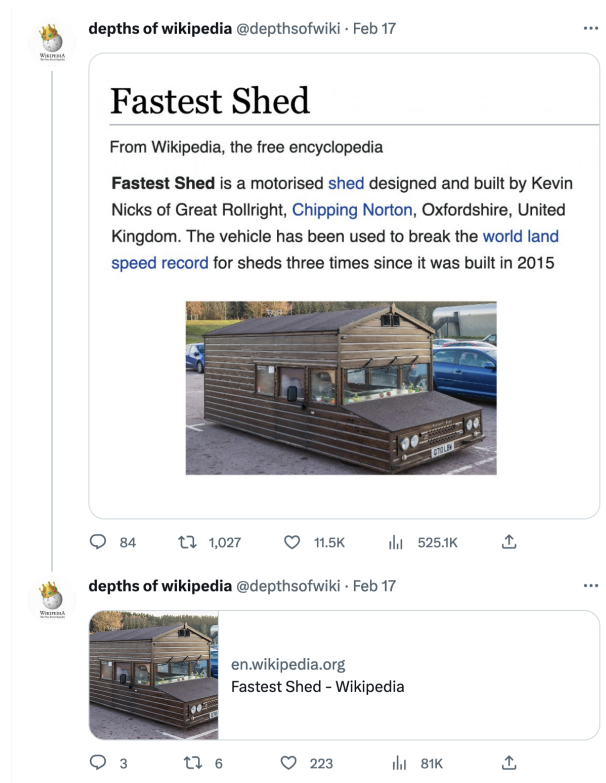


Figure 1: An example of a tweet from depths of wikipedia with linked article.

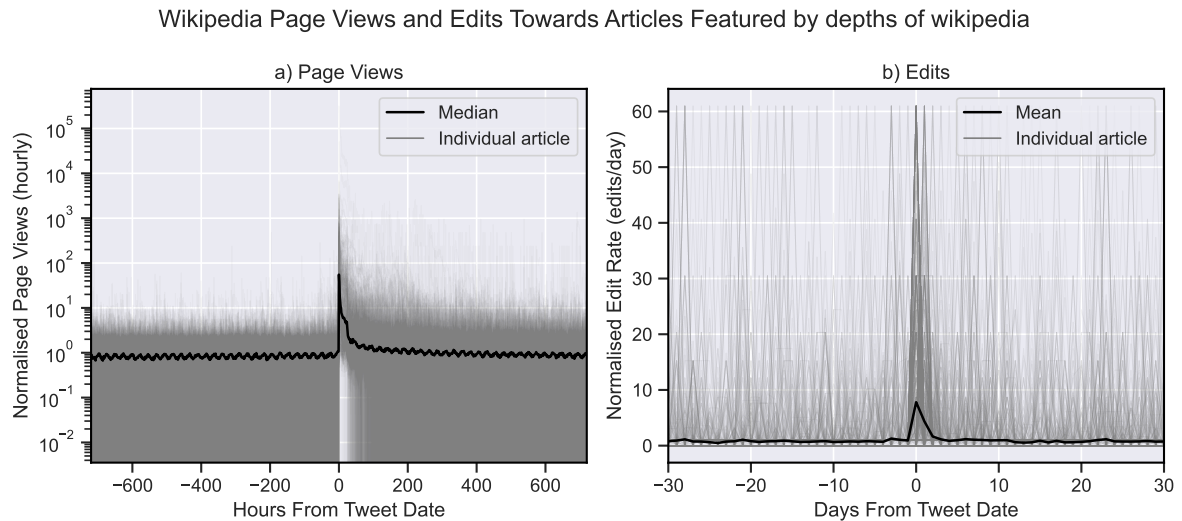


Figure 2: Spikes in both a) page views and b) editing activity on Wikipedia are observed when an article is featured by depths of wikipedia. Page views and edits are normalised to average activity in the 30 days before featuring.