

Investigating Western Bias in Wikipedia Articles about Terrorist Incidents

MARIELLA STEINKASSERER, Graz University of Technology, Austria

THORSTEN RUPRECHTER*, Graz University of Technology, Austria

DENIS HELIC, Graz University of Technology, Austria

Wikipedia is one of the most popular sites on the web and an essential knowledge source for millions of web users. As this online encyclopedia significantly shapes our image of the world, its neutrality is of utmost importance. However, past studies demonstrated that Wikipedia is frequently unbalanced or biased in various subject areas. In this work, we analyze one potential bias on Wikipedia, the so-called western bias, by examining Wikipedia articles about terrorist incidents. To that end, we investigate pageviews of 2029 articles in the first month after events happened. While our findings do not apply to all events in general, we do find that terrorist incidents happening in the “western world” receive more attention on Wikipedia than those from non-western regions.

CCS Concepts: • **Human-centered computing** → **Empirical studies in collaborative and social computing**; **Wikis**.

Additional Key Words and Phrases: Wikipedia, western bias, North-South divide, terrorist incidents, data science

ACM Reference Format:

Mariella Steinkasserer, Thorsten Rupprechter, and Denis Helic. 2021. Investigating Western Bias in Wikipedia Articles about Terrorist Incidents. In *The 17th International Symposium on Open Collaboration (Companion), September 15–17, 2021, Madrid, Spain*, 4 pages.

1 INTRODUCTION

With more than 56 million articles created by millions of contributors since its inception in 2001, Wikipedia is nowadays the most famous online encyclopedia. People use Wikipedia for a variety of reasons, whether it be as a starting point for research, for learning about topics of interest for school or work, or for fact checking [10]. Through this, the encyclopedia is accessed millions of times every day from all over the world, with the most popular English language version being visited around 10 billion times every month.¹

Despite its success, Wikipedia is often criticized. For example, although one of the central basic principles of the website is neutrality, many people accuse it of being unbalanced [4, 5]. Recently, researchers also uncovered inequalities [9, 15] and biases on Wikipedia [1, 7, 12]. In this work, we address a particular kind of bias: *Western bias*, which describes the tendency of western states or regions to regard western behavior patterns as usual, natural, or more important, rendering non-western views as inferior [6]. Past work highlighted such inequalities in other media and concluded that events or locations from some regions of the world receive less attention than others [3, 13].

Depending on the line of research, there are different concepts for dividing the world into western and non-western regions. For our work, we choose the socio-economic division into *Global North* and *Global South*, henceforth referring to western (developed) regions as the *North* and non-western (developing) regions as the *South* [2]. We consider this classification instead of a strict geographical division and investigate whether articles belonging to different regions receive uneven amounts of attention, possibly due to large communities of readers and editors mostly stemming

* Corresponding author.

¹Wikipedia Statistics: <https://stats.wikimedia.org/#/all-wikipedia-projects> and <https://stats.wikimedia.org/#/en.wikipedia.org>

from the North.² These users' interests and media diet shape Wikipedia content quite significantly, possibly omitting perspectives and knowledge other demographic groups could contribute.

In our analysis, we focus on articles about terrorist incidents from 2018 to 2020, which generally attract a lot of attention. We divide these articles according to the economic region the incidents happened in—the South or the North. Through this, we aim to answer the following research question: *Do Wikipedia articles about terrorist incidents in the Global South receive less attention than articles about similar incidents in the Global North?*

Overall, we analyze pageviews of 2029 Wikipedia articles (228 terror, 1229 sports, and 572 election articles), showing that articles about terrorist incidents in the South generally receive less attention than those from the North, seemingly confirming Wikipedia's alleged western bias. However, we also find that this does not apply to other events covered on Wikipedia. For example, elections or sports events show diverging results, with articles covering the South generating as much or even more traction than their northern counterparts. Altogether, our findings further the understanding of western bias on Wikipedia, the mitigation of which would advance neutrality and knowledge equity on Wikipedia.

2 MATERIALS AND METHODS

Approach. To measure the attention an article receives, we collect pageviews within the first four weeks after an event took place. We parse the dates and location from the corresponding Wikipedia article's info box. To compare articles about terrorist activities in the South and North, we apply Mann-Whitney U tests to test for differences in distributions of pageviews and edits. Additionally, we compare results for articles about terrorist incidents to those about other events to check whether such differences are observable regardless of the topic. We use sports events and elections for this purpose, as these can be reliably linked to a location and date. Additionally to pageviews, we also analyze the number of page revisions, as frequent edits could also indicate greater interest in an event. Lastly, we fit a linear regression to assess the effect of the region on the pageviews of terror articles, while also considering the number of deaths.

Dataset. We retrieve article data about terrorist incidents from 2018 to 2020 from lists on Wikipedia [14]. In addition to these lists, we retrieve all articles belonging to the Wikipedia categories *Terrorist_incidents_in_<continent>_in_<year>* to our dataset via *Petscan*.³ Next, we download revision, user, and page history data from MediaWiki history dumps and retrieve the pageviews of the individual articles using the Wikimedia REST API.⁴ Afterward, we assign the respective economic region to each article [8]. Besides the categories containing terrorist incidents, we repeat this process for sports and election categories (*<year>_in_<continent>_sport* and *<year>_elections_in_<continent>*, respectively).

3 RESULTS AND DISCUSSION

Figure 1 (left) shows that the number of pageviews of terrorist incidents from the South is lower than those of articles from the North. We statistically confirm the observed difference between the two groups (Mann-Whitney U, $p < 0.001$). We further test for a difference in the number of edits and conclude that there are significantly fewer edits on articles about terrorist activities from the South ($p < 0.001$). The distribution of pageviews depending on the continent on which the incident took place also reflects this north-south divide (omitted from paper for brevity). Correspondingly, the continental regions that mainly belong to the North (Europe, North America, Oceania) get slightly more pageviews than those from the South (Africa, Asia, South America). Altogether, we conclude that the North's terrorist activities receive more attention than the South's. One possible reason our data shows higher pageviews for the North might be that the western media is mainly concerned with this region, while the South is more of a secondary consideration.

²Wikipedia demographics: <https://en.wikipedia.org/wiki/Wikipedia:Wikipedians#Demographics>

³Petscan: <https://petscan.wmflabs.org>

⁴MediaWiki history dumps: https://dumps.wikimedia.org/other/mediawiki_history/readme.html, Wikimedia REST API: https://wikimedia.org/api/rest_v1

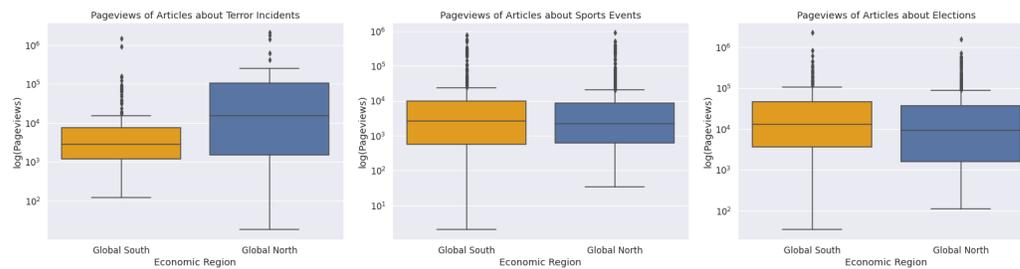


Fig. 1. **Distribution of pageviews of different event types depending on the economic region.** On the left, the events' type is terrorist incidents. In the middle, the pageviews are from articles about sports events and on the right from elections.

This bias might be mirrored by the English Wikipedia, due to its predominantly western user demographic. Moreover, people from the South might use other information channels than Wikipedia to follow such events.

To control for the event type, we also analyze sports and elections. On the one hand, sports events that have taken place in the South do not receive fewer pageviews than those in the North (Mann-Whitney U, $p = 0.374$). However, a comparison of the edits suggests that articles from this category in the South change more often ($p < 0.001$). On the other hand, contrary to our previous findings, articles about elections that took place in the South are getting more attention than those in the North ($p < 0.01$), although the analysis of the distribution of the edits shows no significant differences in the two regions ($p = 0.097$). According to the democracy index published by *The Economist* [11] there are still countries in the South, particularly in Africa, which are not democratic or are currently transitioning into a democracy. This could possibly lead to democratic elections from this region drawing more attention. We leave the exploration to which specific factors (e.g., region or editor characteristics) explain these inequalities for future research.

Finally, we fit a linear regression to determine whether the attention, i.e., pageviews (V), is affected by the economic region (R) while controlling for the number of inflicted deaths (D). To account for outliers, we use log scale for pageviews and deaths. Region is a binary variable coded as 0 for North and 1 for South. To capture the region-specific attention gain a surplus of deaths might bring, we also include the interaction between economic region and deaths. We model the regression as following: $V = \beta_0 + \beta_1 D + \beta_2 R + \beta_3 DR + \epsilon$. The fitted regression (adjusted $R^2 = 0.315$) suggests that deaths is a significant predictor of the attention an article receives ($\beta_1 = 1.811$, $p < 0.001$). Most notably, no significant difference between the North and South exists for when there are no or only few deaths ($\beta_2 = -0.771$, $p = 0.18$). However, the coefficient of the interaction term signals that with an increase in reported deaths, articles in the South receive significantly less attention than those in the North ($\beta_3 = -1.277$, $p < 0.001$). Hence, although our model finds no fundamental regional difference in attention for a terror incident happening, incidents in the North generate far more pageviews as soon as casualties increase—suggesting a bias in terms of coverage of incidents in non-Western countries. We also fit this regression using edits as the dependent variable and find no significant differences to pageview results.

4 CONCLUSION

In this experience report, we demonstrate our approach of showing that Wikipedia articles about tragic events such as terrorist incidents happening in the Global South receive less attention than those in the Global North, manifesting a western bias. However, we also reveal that this does not hold for all other events, such as sports events or elections. Therefore, we cannot confirm the existence of western bias for all articles about events. We further conclude that there are possibly other factors influencing the attention of individual articles and plan to address this in future research.

REFERENCES

- [1] Ewa S. Callahan and Susan C. Herring. 2011. Cultural bias in Wikipedia content on famous persons. *Journal of the American Society for Information Science and Technology* 62 (2011), 1899–1915. <https://doi.org/10.1002/asi.21577>
- [2] Encyclopedia.com. 2021. *North And South, The (Global)*. Encyclopedia.com. Retrieved 2021-07-06 from <https://www.encyclopedia.com/social-sciences/applied-and-social-sciences-magazines/north-and-south-global>
- [3] Mark Graham, Bernie Hogan, Ralph K. Straumann, and Ahmed Medhat. 2014. Uneven Geographies of User-Generated Information: Patterns of Increasing Informational Poverty. *Annals of the Association of American Geographers* 104 (2014), 746–764. <http://www.jstor.org/stable/24537592>
- [4] Federico Guerrini. 2018. *Wikipedia Releases Transparency Report And Pledges To Improve Diversity*. Forbes. Retrieved 2021-05-27 from <https://www.forbes.com/sites/federicoguerrini/2018/07/25/wikipedias-ultimate-challenge-decolonizing-internet-knowledge/?sh=362547a2630b>
- [5] Valerio Lorini, Javier Rando, Diego Sáez-Trumper, and Carlos Castillo. 2020. Uneven Coverage of Natural Disasters in Wikipedia: the Case of Flood. *CoRR abs/2001.08810* (2020), 1–17. <https://arxiv.org/abs/2001.08810>
- [6] Denis McQuail. 2000. Some reflections on the western bias of media theory. *Asian Journal of Communication* 10, 2 (2000), 1–13.
- [7] Felipe Ortega, Jesus M. Gonzalez-Barahona, and Gregorio Robles. 2008. On the Inequality of Contributions to Wikipedia. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*. IEEE, IEEE, <https://ieeexplore.ieee.org/document/4439009>, 304–304. <https://doi.org/10.1109/HICSS.2008.333>
- [8] Wikimedia Foundation Research and Data Analysis. 2021. *Editing-movement-metrics*. Wikimedia Foundation. Retrieved 2021-07-06 from <https://github.com/wikimedia-research/Editing-movement-metrics>
- [9] Aaron Shaw and Eszter Hargittai. 2018. The pipeline of online participation inequalities: The case of Wikipedia editing. *Journal of communication* 68, 1 (2018), 143–168.
- [10] Philipp Singer, Florian Lemmerich, Robert West, Leila Zia, Ellery Wulczyn, Markus Strohmaier, and Jure Leskovec. 2017. Why We Read Wikipedia. *CoRR abs/1702.05379* (2017), 1591–1600. <http://arxiv.org/abs/1702.05379>
- [11] The Economist INTELLIGENCE UNIT. 2020. *Democracy Index 2020: In sickness and in health?* Technical Report. THE ECONOMIST INTELLIGENCE UNIT. https://www.eiu.com/n/campaigns/democracy-index-2020/#mktoForm_anchor
- [12] Claudia Wagner, David García, Mohsen Jadidi, and Markus Strohmaier. 2015. It's a Man's Wikipedia? Assessing Gender Inequality in an Online Encyclopedia. *CoRR abs/1501.06307* (2015), 454–463. <http://arxiv.org/abs/1501.06307>
- [13] Robert West and J. Pfeffer. 2017. Armed Conflicts in Online News: A Multilingual Study. In *ICWSM*. AAAI, <https://ojs.aaai.org/index.php/ICWSM/article/view/14889>, 309–318.
- [14] Wikipedia contributors. 2021. List of terrorist incidents in 2020 — Wikipedia, The Free Encyclopedia. Retrieved 2021-07-06 from https://en.wikipedia.org/w/index.php?title=List_of_terrorist_incidents_in_2020&oldid=1032096613
- [15] Jinhyuk Yun, Sang Hoon Lee, and Hawoong Jeong. 2019. Early onset of structural inequality in the formation of collaborative knowledge in all Wikimedia projects. *Nature human behaviour* 3, 2 (2019), 155–163.