Bias in Wikipedia

Christoph Hube
L3S Research Center, Leibniz University of Hannover
Hannover, Germany
Supervised by Claudia Niederee
hube@L3S.de

ABSTRACT

While studies have shown that Wikipedia articles exhibit quality that is comparable to conventional encyclopedias, research still proves that Wikipedia, overall, is prone to many different types of Neutral Point of View (NPOV) violations that are explicitly or implicitly caused by bias from its editors. Related work focuses on political, cultural and gender bias. We are developing an approach for detecting both explicit and implicit bias in Wikipedia articles and observing its evolution over time. Our approach is based on different factors of bias, with the most important ones being language style, editors, and citations. In this paper we present the approach, methodology and a first analysis.

1. PROBLEM

The *Neutral Point of View* (NPOV) is one of the main principles of Wikipedia. It demands Wikipedia editors to put their personal opinions on a topic aside and create objective content. According to the NPOV policy, within an article, all important opinions on a topic should be represented without any attempt on trying to convince the reader of any of the presented views.

Describing different points of view (POV) is indeed quite uncommon for an encyclopedia and the results of this process do not always convince every editor and reader. While studies have shown that Wikipedia articles exhibit quality that is comparable to conventional encyclopedias [9], research still proves that Wikipedia, overall, is prone to many different types of NPOV violations that are caused by bias resulting from its editors. Research presents different types of bias, such as gender bias [23], cultural bias [1], and political bias [12]. Bias can be explicit, e.g. when statements in an article support a certain POV, or when the cited sources are politically biased [14]. In the case of implicit bias, an article focuses on a specific aspect of the topic but omits other aspects of equal importance.

In some cases, where editors do not find an agreement, the disputes about the content of an article lead to an *edit war*. Two or more opposing sides each struggle to strengthen one POV in the article's content by permanently adding text supporting their own POV or removing text that supports the opponent's POV (typically by reverting the edit). An usual cause of edit wars in Wikipedia articles are controversial topics. Examples are *Global Warming*, *Homosexuality* and *Abortion* [21]. The edit war about the article *Gamergate Controversy*[1] drew public attention after several editors had been sanctioned and even Wikipedia founder Jimmy Wales took part in the discussion. Supporters of the *gamergate movement* state that the Wikipedia article is highly biased towards a pro feministic and pro journalistic POV[2].

In a conflict between two groups of editors A and B it can happen that group A overpowers group B because of a higher number of supporters or because of the supporters being more experienced or technically skilled. As a result, the article is very likely to contain bias towards the POV of group A. Morgan et al. [17] show an example for this kind of situation. In the controversial article *Jyllands-Posten Muhammad Cartoon Controversy*[3] the group of more western oriented editors overpowered another group that tried to remove images taken from the *Muhammad Cartoons* that were depicted in the article. The analysis of Morgan et al. suggests that the English Wikipedia contains a western cultural bias even though it is intended to be a source for people from all nations and cultures. Wagner et al. [23] find that many Wikipedia articles show gender bias which might be caused by the fact that only 15% of Wikipedia editors are female. Wikipedia was also criticized by christian conservatives for showing strong liberal bias. They state that "[a]lthough Wikipedia claims to have credibility because anyone can edit it, in fact the website represents the viewpoint of its most strident and persistent editors."[4]. This accusation even led to the creation of *Conservapedia*[5], a Wiki shaped according to right-conservative ideas.

In this thesis we tackle bias in Wikipedia, and define[6] it as a form of over-representation of one or more POVs compared to all existing POVs, hence, an explicit or implicit violation of the NPOV principle. Is the bias a result of one group of editors overpowering another group (as in the example of

WWW'17 Companion, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4914-7/17/04.
http://dx.doi.org/10.1145/3041021.3053375

[1] https://en.wikipedia.org/wiki/Gamergate_controversy
[2] https://www.theguardian.com/technology/2015/jan/23/wikipedia-bans-editors-from-gender-related-articles-amid-gamergate-controversy
[3] https://en.wikipedia.org/wiki/Jyllands-Posten_Muhammad_cartoons_controversy
[4] http://www.conservapedia.com/Examples_of_Bias_in_Wikipedia
[5] http://www.conservapedia.com
[6] Related work similarly defines the different types of bias.

the *Cartoon Controversy*)? Or did the editor(s) not know about different opinions on the topic (a phenomena known as *"echo chambers"*[7])? Secondly, we observe how bias in Wikipedia evolves. Providing enough time and editors, is bias a recurrent factor or will the content of a Wikipedia article converge to an unbiased content?

Desired properties of our approach are:

1. Determine the factors of implicit/explicit bias
2. Describe the evolution of bias

Given these two problems we postulate the following research questions:

RQ1: What are the factors of bias in Wikipedia?
RQ2: How does bias in Wikipedia evolve?

RQ1 covers the detection of bias using different factors as evidence. Closely related to this question is the question of what causes bias in the first place. We plan to analyze which factors are crucial for the detection of explicit bias and which factors provide evidence of implicit bias.

RQ2 focuses on the evolution of bias over time. We will analyze which parameters (e.g. number of editors), user actions (e.g. deletes) and external events (e.g. elections, scandals) have an influence on the bias of an article.

We believe that bias detection and the neutrality of public knowledge sources such as Wikipedia are crucial for people's opinion forming and the breaking of echo chambers.

2. STATE OF THE ART

In this section we give an overview of related work on different types of bias with a focus on Wikipedia.

2.1 Bias in Wikipedia

Some aspects of bias in Wikipedia and in collective intelligence systems in general have already been studied. Das and Lavoie [3] make use of editor behavior and interactions between editors to determine the topics an editor is interested in and the editor's stance. The assumption is that an editor who reverts the change of another editor, they are likely to disagree. To infer topics and POV of an editor, they use a modified LDA approach in which editors play the role of documents and Wikipedia articles act as words. For evaluation they detect pairs of antagonistic editors which have reported each other. Das and Lavoie also look at the Wikipedia revision history to study the evolution of bias in Wikipedia. They show that bias exists for example where articles on otherwise controversial topics are dominated by a single POV in Wikipedia. This approach differs from our approach in the way that it only focuses on editor interactions and at the topical level, without including other causes of bias, such as the cited sources, the language of editors. We see their work as complementary to ours.

In previous work Das et al [4] analyzed the behavior of Wikipedia admins. They found cases in which Wikipedia editors followed the rules to become admins and shortly after their promotion changed their behavior by forcing their POV on specific topics. These findings are alarming since they show that editors often try to hide their intentions and act objective even though being clearly biased. This reveals another major concern with Wikipedia admins and their influence on the objectivity of an article.

Greenstein and Zhu [12] study political bias in Wikipedia by measuring how much an article is leaning to the liberal or conservative side. They make use of [8], which was initially developed to determine the political slant of newspapers by looking for 1000 specific phrases that were typically used by democratic or republican congress members. Greenstein and Zhu use this approach on articles related to American politics. They find that in its early years Wikipedia on average used to have a liberal bias but that this bias decreased with time. While the number of revisions and editors working on an article usually has a small reducing effect on the bias, the main reason for the decline is the fact that the number of articles has grown so much that the liberal bias plays a less important role on average. The approach in [12] has a clear disadvantage. It relies heavily on the given list of phrases and it is questionable if an old list still has a strong enough meaning to be used as a base for measuring political bias.

2.2 Cultural bias in Wikipedia

The term *cultural bias* is used in different contexts in research about Wikipedia. On one side it refers to differences between Wikipedia language versions, such as the focus of a language version on specific topics or persons [1] or different language versions showing different POVs or different sentiments on the same topic ([25, 26]). The term also refers to the western bias of English Wikipedia on some topics as shown by Morgan et al. [17].

2.3 Gender bias in Wikipedia

Several studies have analyzed the representation of women in Wikipedia compared to that of men ([23, 19, 11]). Wagner et al. [23] do their analysis regarding the dimensions coverage bias (number of articles about notable women compared to that of men), structural bias (links between Wikipedia articles, e.g. do articles about women link articles about men more often than vice versa?), lexical bias (vocabulary used to describe women and men) and visibility bias (How many articles of men and women make it to Wikipedia's front page?). Their results show that women are not underrepresented in Wikipedia but are often presented in a different way, for example an article about a woman is more likely to contain information about the family and romantic relationships. Reagle and Rhue [19] compare gender bias in Wikipedia and Encyclopedia Britannica and conclude that Wikipedia has more and longer articles on women but at the same time articles on women are more likely to be missing in Wikipedia compared to articles on men.

2.4 Political bias in News and Social Media

Lance [14] shows that news media are inherently biased in their reporting. Researchers found that some media outlets (e.g. *The Guardian, New York Times, New Yorker*) are preferred by more liberal thinking persons while other outlets (e.g. *Fox News, Breitbart, The Blaze*) are preferred by more conservative thinking persons[7].

Research shows that the bias of a news outlet can be determined. Dallmann et al [2] use measures for coverage bias (party as main article, mentions of party and party members) and statement bias (sentiment analysis, vocabulary similarity) to determine the political slant of German newspapers towards different German political parties. Zhou et al. [24] leverage user reactions to news articles to classify news as either liberal or conservative.

[7] `http://www.journalism.org/2014/10/21/political-polarization-media-habits/`

Most work on measuring political bias in social media focuses on Twitter and determining the political leaning of Twitter users using their tweet contents and relations to other users ([10],[20],[15]). Makazhanov et al. [16] show that language features can be used as a strong indicator for political leaning. They build language models from all tweets of a political party's candidates and match these with the language models of other Twitter users to determine if these users have a political leaning towards one of the parties. Pennacchiotti et al. [18] also report that language features perform better than social graph features for determining political leanings in Twitter. These findings can not be transferred directly to Wikipedia since language style in Twitter differs from that used in Wikipedia.

3. PROPOSED APPROACH

To answer our two research questions, we analyze possible factors of bias in Wikipedia. The factors presented in this chapter will work as evidence to detect bias (both explicit and implicit) and observe the evolution of it. While our analysis is proceeding, more factors might be added.

We analyze the following factors in the context of RQ1:

1.1 Language Style
1.2 Editors
1.3 Cited Sources
1.4 Gatekeeping and Coverage of Topics

Language Style.

The language used by an editor can be an indicator of its opinion on a topic, and thus leads to explicit bias in an article. For example, it makes a significant difference whether an article refers to a group as "soldiers" or as "terrorists". Language features have proven to be effective when determining the political leaning of a user on Twitter [18]. In Wikipedia, however, tasks like sentiment analysis might be harder since editors tend to use a more neutral style of writing. We are planning to include a phrase based approach analog to Greenstein and Zhu [12] that uses lists of terms and phrases that are typically used by one of the opposing sides in a conflict. The lists have to be up to date and must fit the particular topic. We plan to mine lists of biased terms (e.g. "terrorist") in order to detect biased statements.

Editors.

Das et al. [4] showed that even Wikipedia admins in some cases tend to show strong bias and are forcing their opinion on articles. Can bias be measured by determining the editors that were involved in the creation of an article given that we are able to obtain basic information about these editors, such as the statements that an editor added and removed or which other articles it has worked on? Are specific editors (e.g. novices, veterans, admins, editors working on different topics or editors working on one topic) more likely to add biased content than others? The advantage of analyzing editors over analyzing whole articles is that while articles can contain parts with different bias, for editors we can assume that each editor has only a single POV on a topic. Based on their contributions we plan to create profiles for editors including their interests, expertise, and bias on certain topics. We will also look at how editors cooperate ([13, 3]) and how groups of editors with similar opinions form. This brings us close to the work about echo chambers [7, 22]. In some cases editors might not be aware of their own bias due to a lack of opinion diversity in Wikipedia. We will use the Wikipedia history to extract the text that has been added or deleted by an editor. Through the revisions of all articles in a Wikipedia language version we determine which articles and categories an editor has worked on.

Cited Sources.

Wikipedia demands of its editors to cite sources to prove their statements. Statements without a reference to an accepted source are likely to be removed. An often used source are news media [5] which are proven to be biased [14], e.g. by covering specific topics in more detail than others. It is likely that an editor cites a source that supports his POV. For example some media outlets in the US are known to show liberal and conservative bias. Citing only typically liberal newspapers could show that an editor has a liberal political bias. An example for media bias is the MH17 flight crash in the Ukraine in 2014. While western media outlets like CNN made Russian separatists responsible for attacking the plane, Russian news media such as RT published news stories that should show that the attack was carried out by Ukrainian soldiers [8]. In this thesis, we will focus only on news sources (apart from other cited sources in Wikipedia like web, blogs etc.), due to quality aspects, which tend to be more authoritative than other sources like blogs, etc.

Gatekeeping and Coverage of Topics.

Saez et al. [20] use gatekeeping and coverage as features to measure bias of news outlets and related communities on Twitter. We are planning to use these features to detect implicit bias in Wikipedia. In our case, Gatekeeping refers to the topics that are or are not included in an article. For example the article about *Donald Trump* does not contain the accusation of Trump harassing a teenage girl, even though this topic has been covered by the media[9]. Coverage refers to the extent by which certain topics are covered. An example could be that the controversy section of a biographical article is larger than any other section of the article.

We assume that analyzing the possible factors of bias will give us a better understanding of bias in Wikipedia. Since articles in Wikipedia are dynamic, meaning they can be changed at any time, the bias of an article can also be considered to be dynamic. Changes, like adding, deleting and reverting influence the bias of an article. We want to analyze how the bias of an article evolves over time while editors are shaping the article. Therefore we are analyzing the following aspects of bias evolution in the context of RQ2:

2.1 How does the number of edits change bias?
2.2 How does the number of editors change bias?
2.3 Does a delete or revert improve bias?
2.4 Do specific events change bias?

How does the number of edits change bias?.

The number of edits for an article is a monotonically increasing function over time. As this number increases does the bias decrease? Does it converge to a normal distribution, where the initial revisions of the entity are more biased and with time the bias tends to normalize to a given average?

[8] `https://www.buzzfeed.com/maxseddon/russian-tv-airs-clearly-fake-image-to-claim-ukraine-shot-dow`

[9] `http://nymag.com/thecut/2016/12/trumps-harassment-of-an-18-year-old-girl-on-twitter-led-to-death-threats.html`

How does the number of editors change bias?.

Linus' Law [12] suggests that increasing the number of reviewers will improve the quality of an article. Does this really hold or are too many contributors counterproductive? Does the expertise of the editors matter? How does the diversity of editors, with respect to their expertise based on the categories they are working on influence the bias? Is it counterproductive for the quality of the article if all editors are from the same community and edit only on certain categories (echo chamber)?

Does a delete or revert improve bias?.

Wikipedia editors can delete or revert existing text that was added earlier by another editor. Once we have extracted the deletes a specific editor has performed on an article, we can analyze whether one delete has influenced the bias of an article by comparing the article's versions before and after the delete. We are planning to analyze which types of deletes and reverts have what impact, e.g. deleting a (specific type of) citation, deleting specific terms or phrases.

Do specific events change bias?.

Do real world events that are related to the article change the bias? First we will analyze if an event leads to a peak in the number of edits or deletes on some articles. Then we will check if different types of events (elections, sport scandals, or wars) have different impacts.

While there have been some attempts to measure the bias of an article [3, 12], none of these approaches have included multiple factors of bias. We are planning to analyze different factors of bias in Wikipedia, also including factors that have not been analyzed before, such as the sources that have been cited by an editor. In contrast to most of the related work, our analysis will not focus on complete articles directly but will instead separate text added by different editors to achieve a better understanding on how different editors behave. Furthermore we are not only planning to measure bias but also to observe the evolution of bias over time.

4. METHODOLOGY

While the methodological part is in its early beginning, here, we lay out the basic idea on the directions we are looking into analyzing the different bias factors and bias evolution as part of RQ1 and RQ2.

Firstly, for each Wikipedia article and its revisions we will decompose it in the following atomic parts: (i) section and subsections, (ii) involved editors, (iii) statements added by individual editors, (iv) sources cited by individual editors, (v) editor interactions (e.g. revert), and (vi) editor profiles (in terms of Wikipedia categories). This decomposition will allow us to isolate the different factors we aim to analyze, and propose sound solutions therein. In the following we sketch methodological aspects for the individual components we aim to analyze in Section 3.

Language Style. Following the approach in [12], we aim at mining phrase and word lists, which we will categorize into the different POV labels (liberal, conservative, etc., or any other arbitrary stances on a topic). We aim to do so by analyzing the *editor interactions* for a given article, and mine the phrases and words from the respective statements added by the corresponding editors. Finally, for the task of categorizing the different phrases, we will employ supervised machine learning approaches.

Editors. In this aspect, we will present editors into a *multi-dimensional profile* which contains attributes such as: topics (in terms of Wikipedia categories), cited sources, word representation and the respective categories (as provided by the language features), and editor interactions. Finally, we employ *community detection* techniques and detected communities from the individual editors through their *profiles*.

Cited Sources. Here we focus on categorizing the news sources according to coverage. We aim at detecting the topic coverage and the stance of newspapers towards a specific topic. We use the Wikipedia article title and additionally the statements added by the editors to retrieve the matching news articles. For this we employ the approach proposed in [6]. Next, based on the categorized phrase and word lists, we will categorize the news sources accordingly. In order to have clean and meaningful categorization of news sources, we will extract the facts from such news articles and categorize only those phrases and words which are attributed to the entities (in this case Wikipedia articles) of interest.

In the case of RQ2, here we aim at providing a detailed analysis about the dynamics of bias in Wikipedia. The analysis here is intertwined with the models built in RQ1. At each state of a Wikipedia article based on our models we are able to measure the bias in an article as a means of over-representation of editors, sources etc., of a specific POV. Therefore, between each transition of an article from revision r_1 to r_2 we can measure the difference in terms of bias.

The main aim of this analysis is to provide a general model for bias in Wikipedia, which would make aware Wikipedia admins, and editors in general, about the content they are adding in an article. Furthermore, this could be used as an internal tool and flag suspicious edits which violate the NPOV policy of Wikipedia.

Evaluation.

An open question in this proposal remains the evaluation of the proposed models. This is a hard task since getting a ground truth for bias in Wikipedia is not trivial. One possibility could be to rely on expert evaluation (for smaller datasets) or crowdsourcing (for larger datasets). Breaking the main tasks into subtasks would make them easier to handle. Some of the subtasks would be *Evaluating phrase and word lists*, *Checking if editors have the same POV*, and *Categorizing news sources.*

5. RESULTS

In order to test some aspects of our approach, we performed a first analysis on the Wikipedia article *Malaysia Airlines Flight 17* (MH17)[10] about the flight crash in the Ukraine in 2014. As we mentioned in 3, the incident was perceived differently in western media compared to Russian media. We analyzed how this is reflected in the Wikipedia article, i.e. to which extent both POVs are represented.

The article was created on July 17th 2014 (the day of the incident) and was revisited 6522 times by 1083 different editors (including 75 edits by 15 different bots). 67.4% of the edits were done by the top 10 most active editors [11]. This already shows that a small group of editors has a high influence on the article's content and could possibly shape it according to their POV. The current version of the article

[10] `https://en.wikipedia.org/wiki/MH17`

[11] `https://tools.wmflabs.org/xtools-articleinfo/?article=Malaysia_Airlines_Flight_17&project=en.wikipedia.org`

contains 15 sections with the section on investigations about the cause of the crash being the largest one. There exists also a separate section on Russian media reception.

For our analysis we retrieved the revision history of the article using the Wikimedia API [12]. From the revisions we extracted the text and all citations added by each editor. From our obtained table we removed all editors who added less than three citations. Of the remaining 77 editors 51 cited only non-Russian sources. The other 26 editors cited both Russian and non-Russian sources with only 7 editors citing more than 50% Russian sources. This indicates that the pro-Russian POV is underrepresented in the article. We took a closer look at the editor with the highest percentage of citations to Russian sources (6 out of 7). The editor performed 28 edits on the article trying to add evidence published by the Russian government into the *Cause of the Crash* section of the article. He also changed the phrasing of multiple sentences from an anti-Russian wording to a more neutral or sometimes pro-Russian wording. Most of his edits were reverted by other editors.

6. CONCLUSIONS AND FUTURE WORK

We gave an overview over the topic of bias in Wikipedia, including the state of the art in research, and presented our new approach based on different factors. In section 1 we introduced our two research questions based on the problems of how to detect explicit and implicit bias and how to observe the evolution of bias. In section 3 we discussed different factors of bias in Wikipedia for which we presented possible methods in 4. Our analysis on the article *Malaysia Airlines Flight 17* indicates a pro-western bias. It also shows that citations can be used to identify editors who are actively trying to change the article's content according to their POV.

We plan to concretize the use of the different bias factors presented in this paper and analyze how well they perform when used as evidence for the detection of explicit or implicit bias. Future analysis will be conducted on a bigger set of articles and will also include information about deleted text.

7. REFERENCES

[1] E. S. Callahan and S. C. Herring. Cultural bias in wikipedia content on famous persons. *JASIST*, 62(10), 2011.

[2] A. Dallmann, F. Lemmerich, D. Zoller, and A. Hotho. Media bias in german online newspapers. In *26th Hypertext*, 2015.

[3] S. Das and A. Lavoie. Automated inference of point of view from user interactions in collective intelligence venues. In *ICML*, 2014.

[4] S. Das, A. Lavoie, and M. Magdon-Ismail. Manipulation among the arbiters of collective intelligence: How wikipedia administrators mold public opinion. In *22nd CIKM*. ACM, 2013.

[5] B. Fetahu, A. Anand, and A. Anand. How much is wikipedia lagging behind news? In *Proceedings of the ACM Web Science Conference*, page 28. ACM, 2015.

[6] B. Fetahu, K. Markert, W. Nejdl, and A. Anand. Finding news citations for wikipedia. In *Proceedings of the 25th CIKM*, pages 337–346. ACM, 2016.

[7] R. K. Garrett. Echo chambers online?: Politically motivated selective exposure among internet news users1. *Journal of Computer-Mediated Communication*, 14(2), 2009.

[8] M. Gentzkow and J. M. Shapiro. What drives media slant? evidence from us daily newspapers. *Econometrica*, 78(1), 2010.

[9] J. Giles. Internet encyclopaedias go head to head. *Nature*, 438(7070), 2005.

[10] J. Golbeck and D. Hansen. Computing political preference among twitter followers. In *SIGCHI*. ACM, 2011.

[11] E. Graells-Garrido, M. Lalmas, and F. Menczer. First women, second sex: gender bias in wikipedia. In *26th Hypertext*. ACM, 2015.

[12] S. Greenstein and F. Zhu. Collective intelligence and neutral point of view: the case of wikipedia. Technical report, National Bureau of Economic Research, 2012.

[13] B. Keegan, D. Gergle, and N. Contractor. Hot off the wiki: dynamics, practices, and structures in wikipedia's coverage of the tōhoku catastrophes. In *7th WikiSym*. ACM, 2011.

[14] B. W. Lance. News: The politics of illusion. *White Plains, NY: Longman*, 1996.

[15] H. Lu, J. Caverlee, and W. Niu. Biaswatch: A lightweight system for discovering and tracking topic-sensitive opinion bias in social media. In *24th CIKM*, 2015.

[16] A. Makazhanov, D. Rafiei, and M. Waqar. Predicting political preference of twitter users. *Social Network Analysis and Mining*, 4(1), 2014.

[17] J. T. Morgan, R. M. Mason, and K. Nahon. Negotiating cultural values in social media: A case study from wikipedia. In *45th HICSS*. IEEE, 2012.

[18] M. Pennacchiotti and A.-M. Popescu. Democrats, republicans and starbucks afficionados: user classification in twitter. In *17th SIGKDD*, 2011.

[19] J. Reagle and L. Rhue. Gender bias in wikipedia and britannica. *International Journal of Communication*, 5, 2011.

[20] D. Saez-Trumper, C. Castillo, and M. Lalmas. Social media news communities: gatekeeping, coverage, and statement bias. In *22nd CIKM*. ACM, 2013.

[21] R. Sumi, T. Yasseri, A. Rung, A. Kornai, and J. Kertész. Edit wars in wikipedia. *arXiv preprint arXiv:1107.3689*, 2011.

[22] V. Vydiswaran, C. Zhai, D. Roth, and P. Pirolli. Biastrust: Teaching biased users about controversial topics. In *21st CIKM*. ACM, 2012.

[23] C. Wagner, D. Garcia, M. Jadidi, and M. Strohmaier. It's a man's wikipedia? assessing gender inequality in an online encyclopedia. *arXiv preprint arXiv:1501.06307*, 2015.

[24] D. X. Zhou, P. Resnick, and Q. Mei. Classifying the political leaning of news articles and users from user votes. In *ICWSM*, 2011.

[25] Y. Zhou, A. I. Cristea, and Z. Roberts. Is wikipedia really neutral? a sentiment perspective study of war-related wikipedia articles since 1945. In *29th PACLIC*, 2015.

[26] Y. Zhou, E. Demidova, and A. I. Cristea. Who likes me more?: analysing entity-centric language-specific bias in multilingual wikipedia. In *31st SAC*. ACM, 2016.

[12] `https://www.mediawiki.org/wiki/API:Revisions`