

Omission of Information: Identifying Political Slant via an Analysis of Co-occurring Entities

Jonas Ehrhardt

University of Amsterdam
jonas.m.ehrhardt@gmail.com

Timo Spinde

University of Constance
timo.spinde@uni-konstanz.de

Ali Vardasbi

University of Amsterdam
a.vardasbi@uva.nl

Felix Hamburg

University of Constance
felix.hamburg@uni-konstanz.de

Abstract

Due to the strong impact the news has on society, the detection and analysis of bias within the media are important topics. Most approaches to bias detection focus on linguistic forms of bias or the evaluation and tracing of sources. In this paper, we present an approach that analyzes co-occurrences of entities across articles of different news outlets to indicate a strong but difficult to detect form of bias: bias by omission of information. Specifically, we present and evaluate different methods of identifying entity co-occurrences and then use the best performing method, reference entity detection, to analyze the coverage of nine major US news outlets over one year. We set a low performing but transparent baseline, which is able to identify a news outlet's affiliation towards a political orientation. Our approach employing reference entity selection, i.e., analyzing how often one entity co-occurs with others across a set of documents, yields an F1-score of $F1 = 0.51$ compared to $F1 = 0.20$ of the TF-IDF baseline.

Keywords: media bias; bias by omission; news articles; co-occurrences

1 Introduction

News articles play a significant role in informing the public about current events. Many individuals source their information about those events pre-

dominantly from news articles, making them a major influence in understanding and developing an opinion about the covered topics (Bernhardt, Krasa, & Polborn, 2008; Hamborg, Donnay, & Gipp, 2019; Spinde et al., 2021b)

The coverage by news outlets – such as newspapers, TV-channels, and online platforms – often inherits a particular political slant (Agirdas, 2015; Gentzkow & Shapiro 2010). The slant impacts the general way of reporting and manifests in formulations, source-selection, and perspectives (Dallmann et al., 2015; Gentzkow & Shapiro, 2010; Hamborg et al., 2019). While this provides readers with a constant perspective on covered topics, securing a loyal readership (Gentzkow & Shapiro, 2010), it carries the risk of reinforcing incomprehensive opinions and ideologies (Flaxman, Goel, & Rao, 2016; Garrett, 2009; Iyengar & Hahn, 2009; Munson & Resnick, 2010). With the increasing polarization of news outlets, media bias detection and highlighting have become an increasingly important topic (Spinde et al., 2020a, 2020b, 2020c).

Media bias appears in a variety of manifestations. It ranges from strategic decisions, like the selection of covered events and the selection of sources (Agirdas, 2015; Bourgeois, Rappaz, & Aberer, 2018; Gruenewald, Pizarro, & Chermak, 2009), over the operative placement of news articles and their size allocation within the news outlet (Stempel, 1969; Stovall, 1985), down to the structural level of information within a news article (Park et al., 2009) and the linguistic level of labeling and word-choices (Bhowmick, 2009; Papa-charissi & De Fatima Oliveira, 2008).

Whereas most manifestations of media bias already have been covered, research towards bias by omission of information is rare (Hamborg et al., 2019; Park et al., 2009). “Journalists need to select sources, e.g., press releases, other news articles, or studies, to be used when writing an article. Ultimately, the journalists must decide which information from the sources to be included and which to be excluded from the article to be written. This step is called commission or omission, and likewise affects which perspective is taken on the event” (Hamborg et al., 2019). Bias by omission of information therefore describes the *wanted or unwanted discrepancy of information* between news articles from different news outlets covering the same topic (ibid.). It manifests in the omission of perspectives or sources. Hence, it shows similarities with source or event-selection bias (ibid.).

In the social sciences, the detection of bias by omission of information heavily relies on human judgment for the selection of topics or search terms,

which are manually applied to the examined corpora (Smith et al., 2001; Stemler, 2001; Spinde et al., 2021c). In computer science, there have so far been no specific approaches for the detection of bias by omission of information (Hamborg et al., 2019).

In this paper, we derive a first basic approach for detecting political slant of news outlets, which sets the base for bias by omission of information, using co-occurrences of entities. We seek to identify whether mentions or commentions of specific entities in news outlets relate to the political ideology of the outlet. Therefore, we calculate co-occurrences over all detected entities in a news corpus. In the current version of the system, we excluded headlines, but we will integrate them in future work. By applying statistical analyses, we can depict relations in political slant between compared news outlets. As this approach heavily depends on the selection of co-occurring entity tuples, we apply and evaluate different entity selection methods, including standard procedures from *Natural Language Processing* (NLP) like *Term Frequency – Inverse Document Frequency* (TF-IDF) or *Latent Dirichlet Allocation* (LDA). We evaluate the differences in political slant between news outlet tuples by performing standard χ^2 -tests.

Our dataset¹ consists of 76,221 news articles from nine major news outlets in the US over one year. The average article length is 632 words. The results with a maximum accuracy of 0.6250 indicate the potential of future research, and our method² has the potential of being easily extended and improved or used as a simple baseline in further research.

This paper makes two main contributions. First, we provide a simple and comprehensible method for the estimation of political slant between news outlets based on omission of information. Second, we evaluate different approaches for the selection of co-occurring entities

The remainder is structured as follows. Section 2 gives an overview of the research on the detection of media bias by omission of information. Section 3 introduces the used datasets. Section 4 derives the methodology. Section 5 evaluates the outcomes and presents the findings. Section 6 discusses limitations and chances. Section 7 closes with the conclusion.

1 The dataset can be found in the GitHub repository.

2 <https://github.com/the-banandit/OmissionOrCommission>

2 Related work

Within this section, we summarize approaches for the detection of bias media which can be applied to detect bias by omission of information. Further, we derive the research gap and introduce our approach.

Although no distinctive research has been carried out on this topic, some approaches can be partially adapted for the detection of bias by omission of information. Bourgeois et al. (2018) developed a method to detect and depict selection bias regarding the selection of covered events within a news outlet, using an interaction matrix of news outlets and events recovered from the *Global Database of Events, Language, and Tone* (GDELT). This interaction matrix monitors a wide range of news outlets, annotates and records global events, and their coverage (ibid.). By modeling the decision process between a tuple of events, they can identify an inherent preference structure, which indicates a bias in event selection (ibid.). While this methodology applies to the event selection process, it could also be used in information selection by replacing events with sources or information fragments.

An approach by García-Urbe (2018) identified event selection bias on front pages through a pairwise comparison between different news outlets. By modeling a discrete choice model for the selection of an event tuple, she was able to predict the preference for the omission of events. Although this approach also predicts event selection bias, its application on the information selection process could predict bias by omission of information.

Due to their successful application in previous research, co-occurrences of entities are a suitable variable for detecting media bias. Since they can carry additional information about the author's perspective on a topic, they can outrun simple quantitative approaches that employ single entities.

Gentzkow and Shapiro (2010) detected the political slant of news outlets in the US by comparing co-occurrences of single entities or phrases that were initially published by members of Congress and later echoed by news outlets. As a result, they were able to calculate a slant ranking for each examined news outlet.

Following a similar approach, Groseclose and Milyo (2005) included phrases from policy groups and think tanks in their methodology, defining their own slant ranking of news outlets.

Dallmann et al. (2015) conducted a bias analysis of German online newspapers. They conducted a sentiment analysis, and analyzed co-occurring

ideological terms and party references in online news articles. Their results showed a significant difference in reporting between different political parties.

Within the social sciences, where researchers have studied various forms of media bias for decades, the detection of bias by omission of information has been approached with the manual construction and application of co-occurring search terms and phrases (Smith et al., 2001). This approach requires specific domain knowledge and is neither time efficient, nor can be easily automated. Hence, we see a potential in the automatic detection of bias by omission of information. However, no research on the detection of bias by omission of information has been carried out in the field of computer science (Hamborg et al., 2019).

Therefore, we propose an approach utilizing co-occurrences of entities between articles from different news outlets, to identify and detect political slant of news outlets as a first step towards detecting bias by omission of information. Since this approach heavily relies on the selected entities, we tested different compiling methods, utilizing standard procedures and techniques from NLP. With this approach, we aim to determine *whether co-occurrences are applicable to indicate a news outlets' political slant based on the omission of information*, and hence can be an indicator for bias by omission of information in news articles reporting on the same events.

3 Data

To test our approach, we employed a dataset scraped from the *Common Crawl Project*³. Since there is no comparable research regarding bias by omission of information, we had to rely on related research regarding other forms of media bias to create our dataset. We identified three leading publications that cluster US news outlets into slant groups (Budak, Goel, & Rao, 2016; Flaxman et al., 2016; Groseclose & Milyo, 2005; Spinde et al., 2021a). We interpolated the findings of the leading publications to identify the slant of a news outlet. As a simplification, we translated the numerical values into three distinct slant groups: *liberal*, *center* and *conservative*.

3 <https://commoncrawl.org/>

We scraped a total of 76,221 articles from nine different news outlets equally distributed over the three slant categories using the Common Crawl API (cf. Table 1). The dataset contains all articles from the “news” section of the year 2011. We consider articles from this section equivalent to front-page articles, containing only developments determined important by the editors. We chose 2011 since Common Crawl provides an uninterrupted dataset of articles over the whole year and all news outlets. Additionally, 2011 lies between the publication dates of the reference studies (Budak et al., 2016; Flaxman et al., 2016; Groseclose & Milyo, 2005), including the election campaigns of the 2012 US elections, which were especially ambivalently covered. The average word-count per article is 632 words.

Table 1: Composition of the main dataset, covering news articles from 2011-01-01 until 2011-12-31, in the category “news”

News outlet	Number of articles	Slant group
Cable News Network (CNN)	2,652	center
Chicago Tribune (CTB)	2,843	conservative
Fox News (FXN)	6,508	conservative
Huffington Post (HFP)	14,876	liberal
National Broadcasting Company (NBC)	3,958	center
New York Times (NYT)	11,281	liberal
Reuters (RET)	16,767	center
Washington Post (WPO)	14,814	liberal
Wall Street Journal (WSJ)	2,522	conservative

4 Methodology

After stating the research objective, we present the approaches used to extract co-occurrences of entities, select entities, and test for slant group affiliation. In this paper, we seek to answer to what extent co-occurrences of entities between articles of news outlets are beneficial towards identifying bias by omission of information. Specifically, we seek to define the most suitable method for the selection of entities.

4.1 Selecting entities

As baselines for the entity selection we chose *Term Frequency – Inverse Document Frequency* (TF-IDF) (Spärck-Jones, 1972) and *Latent Dirichlet Analysis* (LDA) (Blei, Ng, & Jordan, 2003), as they are simple standard measures in NLP. For the TF-IDF baseline we divided our dataset into corpora containing only news articles from one news outlet. We calculated a ranking for all occurring terms based on the TF-IDF statistic and extracted the top ranked $n = 9$ terms per corpus. For the LDA baseline we also divided the dataset into corpora only containing news articles from one news outlet. We performed the LDA on those corpora, setting the number of topics to $m = 5$. For each topic in each corpus, we selected the top ranked entity. We accumulated all entities over all topics and corpora in an entity list.

We tested the performance of two approaches on the baselines. The first approach is based on *one constant reference entity*. We derived the top $n = 9$ co-occurring entities within each news outlet's corpus for the reference entity. The reference entity has to carry the potential of being differently covered by news outlets from different political orientation. Therefore, the co-occurring entities would differ for each news outlet, but presumably be more similar for news outlets from the same slant group. We derived the reference entity from a test-dataset, engaging the mentioned requirements. All baselines and our reference entity approach are unsupervised approaches.

The second approach is based on *manual selection*. Therefore, we read 100 random articles from outlets of the outer spectrum of the slant ratings. We then manually selected $n = 9$ entities that carried the potential of being ambivalently covered by different news outlets from different slant groups (cf. Table 2). The selection criteria consisted of being mentioned in at least 30 of the 100 random articles. A personal emphasis was put on the eligibility of the considered entities. Of course, the emphasis was heavily dependent on our prior knowledge of linguistic media bias and the knowledge about the covered topics' outcomes.

4.2 Estimating affiliation towards slant groups

To estimate affiliation towards a slant group, we measured the dependence of the selected entities and a tuple of news outlets. Therefore, we constructed contingency tables over the two dimensions news outlets and quantified co-

occurrences. To measure the entities' dependence on the news outlets, we ran standard χ^2 -tests on the contingency tables.

We calculated the dependence of all news outlet combinations for each entity selection method. Thus, we were able to group the news outlets. We evaluated the performance of our methods by comparing the outcomes of all tuple tests with the slant group affiliation we derived from the literature (cf. Table 2), providing the metrics accuracy, precision, recall, and F1-score.

5 Results

Within this section, we present and evaluate our results. Our best-performing method excels the lowest-performing baseline's accuracy by 150% and F1-score by 154.5%. The results of the different methods for entity selection are summarized in Table 2. The reference entity method scored best with an accuracy of 0.6250 and an F1-score of 0.5090. The weakest performing method is the TF-IDF argmax with an accuracy of 0.2500 and an F1-score of 0.2000.

For each entity selection method, we ran the χ^2 -tests of independence for all combinations of news outlet tuples. As a threshold for rejecting the hypothesis, and hence a slant group affiliation, we chose $p = 0.05$.

Table 2: Results for the different methods of entity extraction; the metrics picture the performance of predicting slant group affiliation.

Entity selection method	Accuracy	Precision	Recall	F1-score
TF-IDF argmax	0.2500	0.1250	0.5000	0.2000
LDA argmax	0.5278	0.5278	0.5370	0.4963
Manual selection 1	0.4444	0.5151	0.5185	0.4375
Manual selection 2	0.5833	0.5000	0.5000	0.4958
Reference entity	0.6250	0.5089	0.5093	0.5090

The results in Table 2 show an increasing performance for more complex methods for entity selection. Simple methods, like *TF-IDF argmax*, utilize only high-level, general features of documents and corpora. Thus, the utilized features show similarities, particularly for large corpora; low performance was both expected and confirmed.

More sophisticated approaches, like the *LDA* or *reference entity* approaches, are of higher complexity and consider more abstract properties of the corpora, like the distribution of topics and their composition. Hence they are able to assess the omission of information between different corpora comparatively.

Especially the *reference entity* approach seems eligible for the detection of bias by omission of information, as it considers the top n co-occurring entities of a reference entity explicitly. It enables a direct comparison between the weighting of entities over different corpora and represents the omission of topics or information. Although the performance with an F1-score of 0.5090 is relatively low, it outruns the baselines. A variation and improvement of the parameters and settings could increase its performance.

The most sophisticated approach of *manual selection* involves human decision-making. The peculiarities of omitted information are perceived subconsciously and subjectively. Therefore, it is highly dependent on the reader's previous experience and sensitivity towards bias terms.

6 Discussion

Since this paper is a first approach for the detection of news outlets' political orientation based on bias by omission of information, it entails multiple constraints. This section discusses limitations in data, methods, and results.

6.1 Data

As there were no existing datasets for the detection of bias by omission of information, we had to compile a new dataset. Therefore, we interpolated the results of three reference publications that evaluated news outlets' political slant. The resulting dataset contains articles published only in 2011, since the reference publications are from 2005 and 2016, and Common Crawl could provide a complete, uninterrupted dataset for all examined news outlets only in 2011. The discrepancy in the publication dates of the reference publications and our dataset could challenge the internal validity of the interpolated biases. However, there is proof that news outlets that inherit a particular bias tend to retain it due to economic reasoning (Gentzkow & Shapiro, 2010). There-

fore, we could assume that there is no shift in the considered news outlets' inherent bias. Additionally, a period of one year of coverage could be considered too short to generate universally valid results. Since the examined type of media bias is based on the comparison of information, the period does not play a significant role, rather the integrity of the dataset. Providing complete coverage over multiple outlets is more important for evaluating and depicting bias by omission of information.

We simplified the numerical ratings of the slant group estimations into three distinct classes. This simplification can be seen as an oversimplification, which could distort the results. Certainly, a numerical representation of bias provides a more accurate indication of the political slant. As the publications we interpolated are measuring different forms of bias, a simplification into distinct abstract groups ensures the usability of the used classifications for our application. The duality of the United States' political landscape supports a division of outlets into the three used categories.

The article number discrepancy between the news outlets is significant for some outlets. As the outlets employ different guidelines for assigning the news articles to the domains, there are differences in the numbers and content of the articles. We tackle the discrepancy in numbers through our method (cf. Section 4).

6.2 Methodology and results

As this paper provides a first approach to detecting news outlets' political orientation based on bias by omission of information, we set our focus when devising methods on transparency and design simplicity. Therefore, we favored statistical tests over deep models.

For the automated approaches, the final number of co-occurrences might have been too little. We did not test the behavior of different numbers of co-occurrences. We applied the argmax criterion for the news outlets. A proper investigation of the number of co-occurrences, combined with improvements in the methodological approaches, could be a valuable extension of this research.

As our approach's performance heavily relies on selected entities for the co-occurrences, the improvement of the selection process would be a valuable extension. In our approach, we employed basic NLP-methods to select entities for further testing. Those approaches did not show high performance, as they did not consider the peculiarities of bias by omission of information.

A valuable extension to those methods would be the utilization of *ideology-* or *bias-terms* (cf. Dallmann et al., 2015). By combining those with reference entities, e.g., extracted with LDA, our approach's performance could be significantly improved.

Our method showed that co-occurrences are applicable for comparatively classifying the political orientation of news outlets. Though this is not a detection of media bias by omission of information, we consider it as a first step towards identifying potentially biased news outlets, and hence a first step towards identifying media bias by omission of information across news outlets.

7 Conclusion

In this paper, we presented a method that is able to depict the affiliation of news outlets towards a political orientation, based on the omission of co-occurring entities. We consider this method as a first step towards the detection of bias by omission of information in news articles reporting on the same event.

We tested our method on a dataset consisting of news coverage over a period of one year. Although the results show a low performance, we exceeded the baselines derived from standard NLP methods by far, i.e., $F1 = 0.20$ of the TF-IDF baseline compared to $F = 0.51$ yielded by our reference entity selection.

Further research can improve the performance by examining the number of co-occurrences. Also, we seek to focus on more comprehensive evaluations, e.g. testing the approach on datasets that strongly diverge in size and number of news outlets. Further, the entity selection method could be extended by an automatic matching with ideology- or bias-terms, automatic entity extraction, or sentiment analysis.

References

- Agirdas, C. (2015). What Drives Media Bias? New Evidence From Recent Newspaper Closures. *Journal of Media Economics*, 28(3), 123–41. [10.1080/08997764.2015.1063499](https://doi.org/10.1080/08997764.2015.1063499)
- Bernhardt, D., Krasa, S., & Polborn, M. (2008). Political Polarization and the Electoral Effects of Media Bias. *Journal of Public Economics*, 92(5–6), 1092–1104. doi: [10.1016/j.jpubeco.2008.01.006](https://doi.org/10.1016/j.jpubeco.2008.01.006)
- Bhowmick, P. K. (2009). Reader Perspective Emotion Analysis in Text through Ensemble Based Multi-Label Classification Framework. *Computer and Information Science*, 2(4). doi: [10.5539/cis.v2n4p64](https://doi.org/10.5539/cis.v2n4p64)
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Bourgeois, D., Rappaz, J., & Aberer, K. (2018). Selection Bias in News Coverage. In *Companion of the The Web Conference 2018 – WWW '18* (pp. 535–543). New York, NY: ACM.
- Budak, C., Goel, S., & Rao, J. M. (2016). Fair and Balanced? Quantifying Media Bias through Crowdsourced Content Analysis. *Public Opinion Quarterly*, 80, 250–271.
- Dallmann, A., Lemmerich, F., Zoller, D., & Hotho, A. (2015). Media Bias in German Online Newspapers. In R. Bergmann, S. Görg, & G. Müller (Eds.), *Proceedings of the 26th ACM Conference on Hypertext & Social Media – HT '15, {CEUR} Workshop Proceedings* (pp. 133–137). New York, NY: ACM Press.
- Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly*, 80(S1): 298–320. doi: [10.1093/poq/nfw006](https://doi.org/10.1093/poq/nfw006)
- Garcca-Uribe, S. (2018). Multidimensional Media Slant: Complementarities in News Reporting by US Newspapers. *SSRN Electronic Journal*. doi: [10.2139/ssrn.3195751](https://doi.org/10.2139/ssrn.3195751)
- Garrett, R. K.. (2009). Echo Chambers Online? Politically Motivated Selective Exposure among Internet News Users. *Journal of Computer-Mediated Communication*, 14(2): 265–285. doi: [10.1111/j.1083-6101.2009.01440.x](https://doi.org/10.1111/j.1083-6101.2009.01440.x)
- Gentzkow, M., & Shapiro, J. M. (2010). What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, 78(1), 35–71. doi: [10.3982/ecta7195](https://doi.org/10.3982/ecta7195)
- Groseclose, T., & Milyo, J. (2005). A Measure of Media Bias. *The Quarterly Journal of Economics*, 120(4), 1191–1237. doi: [10.1162/003355305775097542](https://doi.org/10.1162/003355305775097542)

- Gruenewald, J., Pizarro, J., & Chermak, S. M. (2009). Race, Gender, and the Newsworthiness of Homicide Incidents. *Journal of Criminal Justice*, 37(3), 262–72. doi: [10.1016/j.jcrimjus.2009.04.006](https://doi.org/10.1016/j.jcrimjus.2009.04.006)
- Hamborg, F., Donnay, K., & Gipp, B. (2019). Automated Identification of Media Bias in News Articles: An Interdisciplinary Literature Review. *International Journal on Digital Libraries*, 20(4), 391–415. doi: [10.1007/s00799-018-0261-y](https://doi.org/10.1007/s00799-018-0261-y)
- Iyengar, S., & Hahn, K. S. (2009). Red Media, Blue Media: Evidence of Ideological Selectivity in Media Use. *Journal of Communication*, 59(1), 19–39. doi: [10.1111/j.1460-2466.2008.01402.x](https://doi.org/10.1111/j.1460-2466.2008.01402.x)
- Munson, S. A., & Resnick, P. (2010). Presenting Diverse Political Opinions. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems – CHI '10* (Vol. 3, p. 1457). New York, NY: ACM Press.
- Papacharissi, Z., & De Fatima Oliveira, M. (2008). News Frames Terrorism: A Comparative Analysis of Frames Employed in Terrorism Coverage in U.S. and U.K. Newspapers. *International Journal of Press/Politics*, 13(1), 52–74. doi: [10.1177/1940161207312676](https://doi.org/10.1177/1940161207312676)
- Park, S., Kang, S., Chung, S., & Song, J. (2009). NewsCube. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems – CHI 09* (p. 443). New York, NY: ACM Press.
- Smith, J., McCarthy, J. D., McPhail, C., & Augustyn, B. (2001). From Protest to Agenda Building: Description Bias in Media Coverage of Protest Events in Washington, D.C. *Social Forces*, 79(4), 1397–1423. doi: [10.1353/sof.2001.0053](https://doi.org/10.1353/sof.2001.0053)
- Spärck-Jones, K. (1972). A Statistical Interpretation of Term Specificity and Its Application in Retrieval. *Journal of Documentation*, 28: 11–21.
- Spinde, T., Hamborg, F., Becerra, A., Donnay, K., & Gipp, B. (2020). Enabling News Consumers to View and Understand Biased News Coverage: A Study on the Perception and Visualization of Media Bias. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 2020.
- Spinde, T., Hamborg, F., & Gipp, B. (2020). An Integrated Approach to Detect Media Bias in German News Articles. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020. Virtual Event China* (pp. 505–506). ACM Press. <https://doi.org/10.1145/3383583.3398585>
- Spinde, T., Lada, R., Mitrović, J., Hamborg, F., Granitzer, M., Gipp, B., & Donnay, K. (2021). Automated identification of bias inducing words in news articles using linguistic and context-oriented features. *Information Processing & Management*, 58(3). <https://doi.org/10.1016/j.ipm.2021.102505>
- Spinde, T., Rudnitskaia, L., Hamborg, F., & Bela, G. (2021). Identification of Biased Terms in News Articles by Comparison of Outlet-specific Word Embeddings. In

Proceedings of the 16th International Conference (iConference 2021). Virtual Event China. Springer Nature [Preprint].

Spinde, T., Rudnitckaia, L., Sinha, K., Hamborg, F., Bela, G., & Donnay, K. (2021). MBIC – A Media Bias Annotation Dataset Including Annotator Characteristics. In *Proceedings of the 16th International Conference (iConference 2021). Virtual Event China.* Springer Nature [Preprint].

Stemler, S (2001). An Overview of Content Analysis. *Practical Assessment, Research and Evaluation*, 7(17), 2000–2001. doi: [10.1362/146934703771910080](https://doi.org/10.1362/146934703771910080)

Stempel, G. H. (1969). The Prestige Press Meets the Third-Party Challenge. *Journalism Quarterly*, 46(4), 699–706. doi: [10.1177/107769906904600402](https://doi.org/10.1177/107769906904600402)

Stovall, J. G. (1985). The Third-Party Challenge of 1980: News Coverage of the Presidential Candidates. *Journalism Quarterly*, 62(2), 266–271. doi: [10.1177/107769908506200206](https://doi.org/10.1177/107769908506200206)

In: T. Schmidt, C. Wolff (Eds.): Information between Data and Knowledge. Information Science and its Neighbors from Data Science to Digital Humanities. Proceedings of the 16th International Symposium of Information Science (ISI 2021), Regensburg, Germany, 8th–10th March 2021. Glückstadt: Verlag Werner Hülsbusch, pp. 80–93. DOI: doi.org/10.5283/epub.44939.