Using NLP to Analyze User and News Bias in the 2023-2024 Hamas-Israel Conflict

Avi Rosenfeld[1]

[1]Jerusalem College of Technology
Department of Computer Science
21 Havaad Haleumi Street, Jerusalem, Israel 9116001
rosenfa@jct.ac.il

Abstract

The Israel-Hamas conflict that began in 2023 has resulted in extensive casualties and destruction, sparking worldwide discussions and garnering significant media attention. Users' and news groups' opinions regarding the conflict on social media underscore strong opinions and biases. Distilling the facts of the conflict from biased opinions is difficult, but we believe linguistic approaches like Natural Language Processing (NLP) can help. We posit that NLP techniques such as unigram word analysis and predictive text modeling can objectively analyze user opinion and media bias. To support this claim, we created accurate predictive models to classify media and user biases with over 0.9 AUC in two datasets based on thousands of user and newsgroup tweets about the conflict. Significant unigram findings include limited discussion among pro-Israel about Palestinian casualties and pro-Palestinian groups making little reference to the actions of Hamas. This study not only provides a methodology applied to this conflict but also serves as a use-case for how NLP can quantify user and news bias, thus allowing people to better objectively evaluate issues with differing opinions.

Introduction

The Israel-Hamas war erupted on October 7, 2023, after Hamas militants stormed into Israel, killing and taking hostage hundreds of innocent civilians. The war and its coverage have sparked intense debate and controversy worldwide. As of the writing of this paper in March 2024, the ongoing conflict has resulted in more than 30,000 deaths and widespread destruction. The war has also exposed many opinions on social media with diametrically opposing views. It has also highlighted the challenges and pitfalls of media coverage in the age of social media, with not all media coverage being equally accurate, objective, and fair. User and media bias intertwine, significantly affecting public opinion, policy making, and conflict resolution while shaping the perception and understanding of reality on the ground (Baum and Potter 2008). Therefore, it is essential to assess and evaluate both public opinion and media bias in the Hamas-Israel war and to create a general methodology for its quantification and evaluation.

However, assessing public opinion and media bias is not a simple or straightforward task. Media bias is a complex and multidimensional phenomenon, which can manifest in different ways and levels, and can vary depending on the topic, the outlet, the audience, and the time (Covert and Wasburn 2007; Ribeiro et al. 2018; Morstatter et al. 2018; de Lima-Santos and Ceron 2021). Moreover, bias is often inherently subjective with people having different opinions and expectations about what constitutes unbiased and balanced coverage. Therefore, traditional methods of assessing media bias, such as manual content analysis or surveys, may be limited, biased, or impractical.

Consequently, we suggest using applied linguistics, particularly Natural Language Processing (NLP) techniques, to identify and then fight bias. NLP has been generally suggested as a powerful tool for analyzing news and media perceptions (Ribeiro et al. 2018; Morstatter et al. 2018; de Lima-Santos and Ceron 2021). It was previously suggested that NLP might be generally used to find implicit biases of online posts, or the "framing" of the information presented (Morstatter et al. 2018). This paper describes how we used NLP to provide insights about these implicit biases about posts regarding the Hamas-Israel War of 2023-2024. We believe these tools help to overcome some of the challenges and limitations of traditional methods, such as scalability, objectivity, transparency, and reproducibility. We believe that the methods provided can generally apply to all conflicts and issues with opposing opinions. Our approach provides a more comprehensive and nuanced understanding of the conflict, revealing patterns, trends, and anomalies that may otherwise go unnoticed. We believe that this in turn enables users to discuss the conflict in a more constructive, objective light.

Our first contribution is the creation of two unique datasets that capture various aspects of social media tweets from the conflict. We used a commercial software package to obtain thousands of posts from X (formerly Twitter) over a month-long period of the conflict. The first datasets focused on capturing user opinion, downloading thousands of posts containing six different trending hashtags used by both sides (three for each side). We also created a second dataset with social media posts by many news outlets. To highlight our methods' objectivity and ensure reproducibility, we released these datasets to the public so our results can be reproduced and further analyzed.

Our second major contribution is using unigram NLP

methods to quantify different biases within the conflict. In general, we found that this relatively simple and transparent analysis was sufficient to point to significant differences of opinion and bias. To quantify these issues, we constructed logistic regression models for both datasets to predict if a tweet was pro-Israel or pro-Palestinian. All models used a unigram analysis with over 0.9 AUC.

We present results highlighting which unigrams underscore a user or media orientation without bias. For example, differences in whether Hamas had *fighters, activists, terrorists* or no mention at all about their actions on October 7th all indicate bias. Although Hamas is widely considered a terrorist group by many countries, we found that pro-Palestinian outlets rarely referred to their actions, and instead focused on the suffering of non-combatant Palestinians. Even the name of this conflict differs in the media. What is often referred to as the "Hamas-Israel" war in Western outlets like CNN and Sky News is referred to as the "War on Gaza" by Al Jazeera. Our thesis is that this difference is intentional and not trivial, and it highlights their implicit bias that this war is not a fight against terrorists but against all Gazan citizens. Similarly, the term *Hamas* is much less present in Al Jazeera than in most Western news outlets. The term *genocide* was only used by pro-Palestinian users and news outlets, except when pro-Israel users stressed their view that no genocide existed. We found that while pro-Israel users often referenced the civilian *hostages* taken on October 7th, pro-Palestinian users typically ignored referencing this term, while focusing on Palestinian civilians killed, particularly those in *hospitals*. While NLP methods cannot determine causality, it has already been suggested by the Palestinian Center for Policy and Survey Research (PCPSR) that only 7% of Palestinians are aware of Hamas's actions on October 7th due to the lack of media reporting. We further discuss these results and suggest how NLP can provide a more balanced perspective of the conflict.

Methodology

To study user and media bias, we downloaded thousands of X tweets over the first month of the conflict (October-November 2023). While various tools exist for downloading tweets, we choose the Octoparse package [1] for two major reasons. First, this package was recognized for its ease of use within previous Twitter studies (Ahamad, Mahmoud, and Akhtar 2017). Second, we utilized Octoparse's built-in templates for downloading the relevant tweets; the straightforward use of these existing templates was a significant consideration in choosing this package.

Our study intentionally focused on English tweets and not those in other languages (e.g. Arabic or Hebrew) for several reasons. Firstly, comparing texts in a single language allows for direct comparison without the need for interpretive translations. Secondly, analyzing text in Semitic languages like Hebrew and Arabic is more intricate due to their complex morphology. Thirdly, English texts are often tailored for a global audience and may exhibit a more moderate tone than texts intended for local readerships. Last, certain packages, like those used in the sentiment analysis, do not support all languages (like Hebrew). However, we believe that a similar study comparing English and non-English tweets in other languages is feasible with our methodology and may yield new insights.

Two datasets were constructed. We first used an Octoparse template to download content from six different commonly used hashtags. These hashtags were chosen as they were identified by the Twitter API as currently trending. Three of these hashtags (#FreePalestine, #IsraelTerrorists, #Palestine_Genocide) were strongly perceived as pro-Palestinian, while three of these hashtags (#HamasIsis, #HamasTerrorists, #StandWithIsrael) were strongly pro-Israel. This dataset was helpful in understanding user's biases within these targeted groups. Second, we analyzed news tweets from a wide range of newsgroups by using Octoparse's template to download content by different X Handles (user names). We selected three strongly pro-Palestinian news handles: @QudsNen, @Ajplus, and @AJEnglish. We also selected three strongly pro-Israel accounts: @TimesofIsrael, @Jerusalem_Post, and @Haaretz-com. Last, we studied a variety of news organization posts taken from @SkyNews, @BBCworld, @CNN, @RT_com, @Telegraph, @FoxNews, and @Independent. Table 1 summarizes the news tweet sources.

Source	**Total**	**Relevant**	**Relevant %**
AJEnglish	540	423	78.3
AJPlus	153	144	94.1
Al Quds	296	294	99.3
Jerusalem Post	339	217	64.0
Haaretz	194	151	77.8
Times of Israel	207	122	58.9
BBCWorld	2261	435	19.2
CNN	509	148	29.1
Foxnews	569	52	9.1
Independent	399	13	3.3
RT.com	536	125	23.3
Sky News	400	65	16.3

Table 1: Summary of news tweet data. The percentages of relevant tweets are given based on their absolute value and the percentage of the total number of captured tweets.

While various NLP methods exist, we intentionally focused on Bag of Word Models (BoW) to analyze these datasets. Despite the advent of more sophisticated models like Large Language Models (LLMs), such as BERT and GPT, several advantages exist for using BoW models. One notable advantage lies in BoW's simplicity and interpretability, making it particularly suitable for tasks where transparency and explainability are paramount. In our datasets this allowed us to highlight the common word frequency differences and through visualization tools such as word clouds. This approach allows us to not only create models that focus on document classification, but also provided a straightforward and efficient representation of which terms were significant in the textual data. While LLMs showcase remarkable performance in capturing contextual nuances

[1] https://www.octoparse.com/

and semantic relationships, their lack of explainability and transparency makes it difficult to assess why these models are more accurate and what terms and phrases are being used that better quantified these biases (Rosenfeld and Richardson 2019).

To facilitate future analysis and reproducibility of our work, we have left both the original dataset without any filtering and the one used within our study at: (removed for blind review, but an anonymous copy of the files exists at: https://tinyurl.com/5n8dxn8u). Note that this directory contains two sub-directories of the source files: unprocessed_files and paper_files. The unprocessed files are the original files, containing a total of 1495 hashtag tweets and 6584 total news tweets. The filtered paper files datasets removed duplicate records, messages without text (graphics only), and foreign language tweets. Furthermore, we balanced the datasets and only kept those tweets with the highest number of views. This left 900 total hashtag tweets (450 for each side), and 2362 relevant news tweets. Relevant news tweets were defined as tweets with the words *Palestinian, Palestine, Israel, Israeli, Gaza, Hamas, Hostage, Antisemitic* or *Islamophobia*. This was done to remove any news posts unrelated to the conflict in our analysis.

Experimental Results

This section presents three types of results. First, we focus on what percentage of newsgroup tweets contained keywords relating to the conflict, showing which groups are more focused on this conflict. Second, we studied the most frequent terms in the tweets across hashtags and news tweets. We present word clouds visually summarizing the key differences between these groups. Last, we focus on studying and explaining which news tweets are most similar to pro-Israel and pro-Palestinian groups.

Referring back to Table 1, note that the percentage of tweets related to the conflict is nearly 100% for Al Quds and AJPlus down to only 3.3% for the Independent. Interestingly, all three Israeli news outlets had lower percentages of relevant tweets (77.8%, 64% and 58.9%) than the Qatari-based Al Jazeera news outlet (78.3%). CNN had the highest relevance score among Western news accounts. This is more than triple than that of the further right-leaning Foxnews, another American-based news outlet (29.1% versus 9.1%). This metric does not allow us to conclude why these outlets choose different stresses of the conflict. There seems to be a correlation between political orientation (e.g., anti-Western and left-leaning) and conflict coverage, almost to an obsessive level. In the case of the Qatari outlets Al Jazeera and AJ+, it appears to leave no other issues as newsworthy other than the conflict itself. We further explore this point below in the discussion section.

The second step of our research produced word clouds for each of the datasets. Each cloud included the top 100 ranked words by frequency for each corpus. The clouds were created using the Orange data mining package (Demšar et al. 2013). Stopwords were filtered out, and lemmatization was done using Orange's tool. These word clouds are found in Figures 1 and 2.

We believe several points are worth noting in these word clouds. We first noted much overlap between the keywords in two sides by examining the set of keyword clouds. Certain words, such as *people*, *Gaza*, and *Israel* have similar term frequencies. Conversely, we noted that the terms *genocide* and *ceasefire* were missing from the pro-Israeli hashtags while the words *Hamas, hostages, terrorist, kidnap* and *October 7th* were much more prevalent in the Israel dataset (Figure 1). Similarly, the most frequent words in the pro-Palestinian tweet word clouds (Figure 2), are *#Palestine Palestine*, *Palestinian*, *hospital*, *bombardment*, *airstrike*, and *genocide*. They terms appear on the Israeli side, albeit less frequently. In contrast, the terms *Hamas*, *hostage*, *deal*, *release*, and *October 7th* are much more common in the pro-Israel news. Overall, these results stress that the pro-Israel side is focused on Hamas's kidnapping hostages on October 7th, and their desire for a hostage deal to gain their release, while the Palestinian side is focused on a ceasefire without any mention of hostages, Hamas or October 7th, to prevent a genocide which focused on airstrikes and attacks, including those on hospitals.

We used Weka's implementations of Multinomial Naive Bayes, Logistic Regression, and Random Forest for prediction models to demonstrate that accurate models to be created based on these differences. Training and testing used the default settings for these models with 10-fold cross-validation. The BoW was created using Weka's StringToWordVector package with default parameters. To prevent overfitting, we removed all hashtag terms from the this dataset as they defined the post as pro-Israel or pro-Palestinian (e.g. #FreePalestine, #IsraelTerrorists, #Palestine_Genocide, #HamasIsis, #HamasTerrorists, #StandWithIsrael). Both prediction models successfully differentiated between the two classes with Area Under the Curve (AUC) above 0.9 for all models and algorithms. The results are found in Table 2.

summary)

Dataset	Algorithm	Accuracy	AUC
Hashtags	Logistic Regression	85.2	0.92
	Naïve Bayes	85.9	0.94
	Random Forest	84.0	0.92
News	Logistic Regression	89.0	0.96
	Naïve Bayes	89.0	0.97
	Random Forest	83.6	0.92

Table 2: Accuracy and AUC (ROC) for the two datasets and algorithms

As the models in the previous section had clear indications of a pro-Israel or pro-Palestine bias, we reasoned that we would similarly be able to test for media bias based on these models. To test this hypothesis, we took the news model built with the assumed Pro-Palestine relevant tweets from AJEnglish, AJPlus, and Al Quds, and compared them with the Jerusalem Post, Haaretz and Times of Israel tweets. We then used a logistic regression model to determine which relevant tweets were classified as pro-Israel or pro-Palestinian. This was done by querying the model to de-

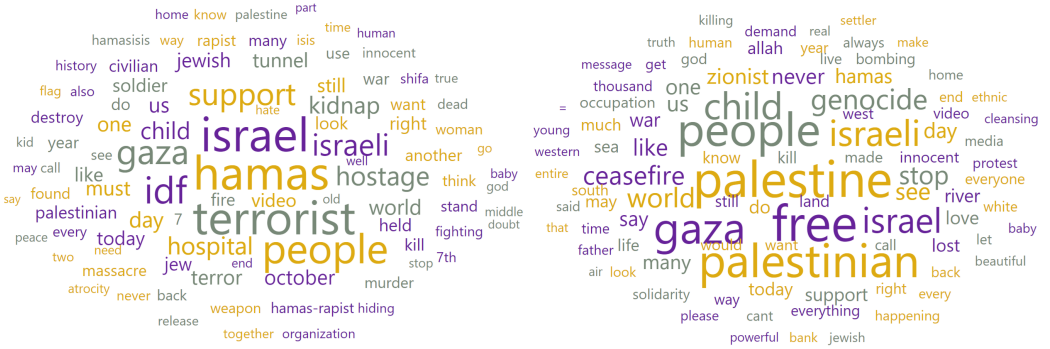

Figure 1: Word cloud of the 100 most frequent words in the hashtag dataset – Israel (left) and Palestinian (right)

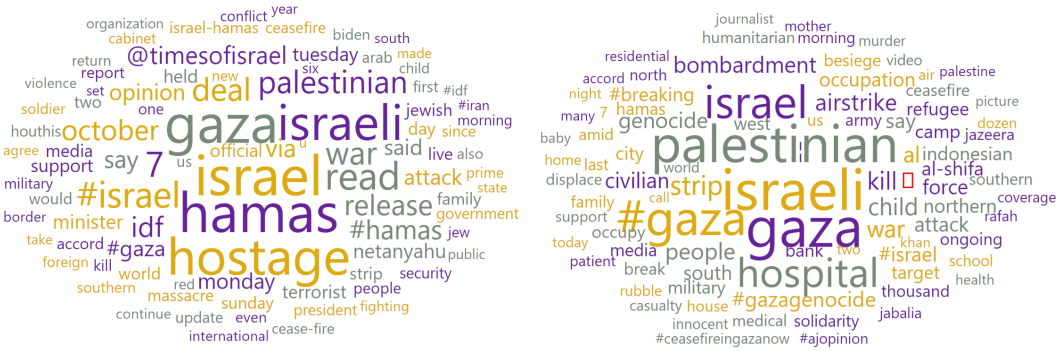

Figure 2: Word cloud of the 100 most frequent words in the news tweets dataset – pro-Israel (left) and pro-Palestinian (right)

termine to which group each tweet was more similar. Note from the results in Table 3 that we did not run this analysis for the Independent as only thirteen tweets were relevant, and thus, the study did not have a statistically significant sample size. Also note that only the BBC and RT (Russian Times) were deemed by the model to be classified more similarly to the pro-Palestinian tweets (58.6% and 61.6% accordingly). All other news groups were more similar to the pro-Israel news with Sky News showing a near 90% of its 65 relevant tweets as being pro-Israel.

Source	Total	Relevant		Pro	
		Count	%	Israel	Palestinian
BBCWorld	2261	435	19.2	41.4	58.6
CNN	509	148	29.1	58.8	41.2
Foxnews	569	52	9.1	71.2	28.8
Independent	399	13	3.3	N/A	N/A
RT.com	536	125	23.3	38.4	61.6
Sky News	400	65	16.3	89.2	10.8
Telegraph	386	36	9.3	63.9	36.1

Table 3: A logistic regression model judged how many tweets from each channel were similar to those of pro-Israel or pro-Palestinian news outlets.

Discussion and Conclusions

We believe that the unigrams used by the different groups belie their biases. For example, the term *genocide* was prominent in the keyword and hashtag dataset only on the pro-Palestinian side. This is consistent with South Africa's contention in the International Court of Justice on January 26th, 2024 that Israel committed genocide during this conflict. While the court did not find that Israel committed genocide, it did require Israel to prevent and punish any public incitement to commit genocide against Palestinians in Gaza. Interestingly, our NLP analysis shows no such incitement or mention of genocide in the Israeli hashtag data. Similarly, this issue did not appear in the collected news data, even within pro-Palestinian outlets [2]. Thus, our NLP analysis shows no evidence of genocide reported in the news or any incitement to genocide in the social media posts.

Similarly, the taking of Israeli citizens as hostages was widely condemned in Western news outlets as in Israel. This issue did not appear with statistical significance in pro-Palestinian news outlets or in their supports on social media. There may be a causal relationship between the lack of reporting by these news outlets and popular perceptions. In one survey conducted by the Palestinian Center for Policy and Survey Research (PCPSR) taken between 22 November and 2 December 2023[3] (close to the time this data was collected), only 7% said that they believe Hamas committed the atrocities on October 7th. This is utterly surprising considering the thorough confirmation Western outlets have pro-

[2]Note that our data collection ended nearly 2 months before the court case.

[3]https://pcpsr.org/en/node/963

vided to Israeli claims about these atrocities. Khalil Shikaki, a professor of political science in Ramallah and director of the PCPSR, who conducted the poll, believes this low value was caused by the lack of coverage in Palestinian and Arab media. While our study cannot judge Al Jazeera, AJ+ or Al Quds's motivation for not reporting about Hamas or their atrocities, it seems clear from our NLP study that this knowledge is not reaching pro-Palestinian users or news media in line with the PCPSR poll. In fact, it is important to stress that Al Jazeera does not refer to this situation as a conflict between Israel and Hamas, but instead as a war between Israel and Gaza [4]. This difference is likely significant – removing Hamas as a central focus for anything that occurred on October 7th or afterward and instead implying that people in Gaza are the target of Israel's war.

When addressing any significance to this or other terms, it is very important to understand the limitations of NLP analysis. The lack of the keyword *genocide* in pro-Israel posts is not conclusive proof that no incitement to genocide existed, only that no statistical evidence exists for the keyword's usage. Similarly, its existence in pro-Palestinian posts does not indicate proof that Israel committed this crime. Similarly, the Palestinian news covered Israel's military action near hospitals. The Israeli news made no significant mention of this keyword. This again shows diverging narratives– while the Palestinian news attributed significance, and typically malice, to Israel's actions, the hospital issue did not appear with a statistically significant frequency on the Israeli side. Again, this difference does not mean Israel did or did not act with malice towards Palestinians in or near hospitals– only that NLP differences exist– which can then be used to stress different sides to the conflict to fight bias.

Conversely, it is important to stress that Israeli opinion and news coverage do not focus on Palestinian civilian suffering. Terms such as *Palestinian, Hospital* are not often found in pro-Israel users or news tweets. This is particularly interesting as we intentionally chose Haaretz as one of the 'pro-Israel' groups. Haaretz is often associated with left wing politics, holding progressive views about Palestinian rights. While NLP cannot provide insight as to why this lack of focus on civilian suffering occurs, other Western outlets, such as LeMonde, have attributed this to an Israeli society that was traumatized by civilian death and hostage-taking on October 7th, and as such is not interested in hearing about Palestinian suffering [5]. Others feel, as freed hostage Mia Schem who spoke of her captivity in Gaza, that "Everyone there is a terrorist"[6]. This sentiment is backed by several polls, including one reported by the Associated Press that over 90% of Gazans support Hamas's actions[7], which blurs the line between an "innocent" civilian and one who "guilty" of supporting Hamas's tactics. However, as previously noted, it is possible that this high support may be due to the biased reporting in pro-Palestinian sources which ignored Hamas' crimes.

The BBC was lambasted in Israel and abroad during this conflict as having an anti-Israeli bias. In fact, Sky News brought as "support" of this bias an episode of the Israeli comedy show *Eretz Nehederet*, acting out a satirical report by the BBC which wrongly claimed that Israel intentionally hit hospitals when they didn't[8]. However, other UK watchers claimed the BBC was biased for Israel. In fact, the UK group More in Common actually found that the BBC was the most balanced of all polled news organizations [9]. Once again, it is important to note the limitations of NLP in judging the BBC or any other news organization. However, our logistic regression model showed that the BBC used language more similar to the pro-Palestinian side of the conflict than that of the pro-Israel position. However, we did check other prediction models and found that a Random Forest model actually showed that 50% of the relevant tweets were pro-Israel and 50% were pro-Palestinian. It is also important to note that being unbiased here is not necessarily desirable. Hamas has been declared by many governments as a terrorist group. No other Western news media was nearly as pro-Palestinian (and by our models this includes ignoring Hamas). Only the Russian news scored similarly.

We recognize that the authors of this paper themselves may have biases that impacted how they presented their results, especially in this discussion section. Nonetheless, our goal is two-fold: First, we believe that this and future social data studies should provide the raw data that formed the analysis basis so other researchers can replicate their results and confirm that their NLP was done without bias. The advantage to this methodology is that transparent NLP is used as the term frequency is not something that can be changed based on a person's bias. As such, we believe that transparent methods, such as term frequency, should be the basis of any applied linguistic NLP analysis.

Finally, it is important to discuss the implications of this paper for future conflicts and user opinion mining. While term mining is the most basic of NLP methods, it proved to be extremely effective in quantifying key differences in the datasets we studied. As such, we believe that it should be equally useful in future studies. Similarly, term analysis has the potential for showing bias within news reporting by identifying which terms are, and equally important, are not, being used. Term analysis is one form of feature analysis that provides models that are inherently more transparent and understandable to people (Rosenfeld and Richardson 2019) and, as such, should be preferred to many types of NLP currently being considered.

[4]https://www.aljazeera.com/tag/israel-war-on-gaza/

[5]https://www.lemonde.fr/en/international/article/2024/01/21/israeli-society-deaf-to-palestinian-suffering-faces-enduring-collective-trauma_6453511_4.html

[6]https://www.timesofisrael.com/liveblog_entry/released-hostage-mia-schem-everyone-in-gaza-is-a-terrorist-i-experienced-hell/

[7]https://apnews.com/article/israel-hamas-palestinians-opinion-poll-wartime-views-a0baade915619cd070b5393844bc4514

[8]https://www.youtube.com/watch?v=GDjs73DND6c&t=9s and other URLs. This viral video has over 1 million views.

[9]https://www.moreincommon.org.uk/our-work/research/more-than-choosing-sides/

References

Ahamad, D.; Mahmoud, D.; and Akhtar, M. 2017. Strategy and implementation of web mining tools. *International Journal of Innovative Research in Advanced Engineering*, 4(12): 1–7.

Baum, M. A.; and Potter, P. B. 2008. The relationships between mass media, public opinion, and foreign policy: Toward a theoretical synthesis. *Annu. Rev. Polit. Sci.*, 11: 39–65.

Covert, T. J. A.; and Wasburn, P. C. 2007. Measuring media bias: A content analysis of Time and Newsweek coverage of domestic social issues, 1975–2000. *Social science quarterly*, 88(3): 690–706.

de Lima-Santos, M.-F.; and Ceron, W. 2021. Artificial intelligence in news media: current perceptions and future outlook. *Journalism and media*, 3(1): 13–26.

Demšar, J.; Curk, T.; Erjavec, A.; Gorup, C.; Hocevar, T.; Milutinovic, M.; Možina, M.; Polajnar, M.; Toplak, M.; Staric, A.; Stajdohar, M.; Umek, L.; Pretnar Žagar, A.; Žbontar, J.; Žitnik, M.; and Zupan, B. 2013. Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research*, 14: 2349–2353.

Morstatter, F.; Wu, L.; Yavanoglu, U.; Corman, S. R.; and Liu, H. 2018. Identifying framing bias in online news. *ACM Transactions on Social Computing*, 1(2): 1–18.

Ribeiro, F.; Henrique, L.; Benevenuto, F.; Chakraborty, A.; Kulshrestha, J.; Babaei, M.; and Gummadi, K. 2018. Media bias monitor: Quantifying biases of social media news outlets at large-scale. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 12, 290–299.

Rosenfeld, A.; and Richardson, A. 2019. Explainability in human–agent systems. *Autonomous Agents and Multi-Agent Systems*, 33: 673–705.