

Human Interest Framing across Cultures: A Case Study on Climate Change

Anonymous ACL submission

Abstract

Human Interest (HI) framing is a narrative strategy that injects news stories with a relatable, emotional angle and a human face, to engage the audience. In this study we investigate the use of HI framing across different cultures in news articles about climate change. Despite having a high impact in the public’s behaviour and perception of an issue, HI framing has been under-explored in NLP to date. We perform a systematic analysis of HI-stories across cultures to understand its role in climate change reporting. Our findings reveal key differences in *how* climate change is portrayed across countries, encompassing aspects such as narrative roles, article polarity, pronoun prevalence, and topics. We also demonstrate that these linguistic aspects boost the performance of fine-tuned transformers on HI story classifications.¹

1 Introduction

News not only summarises the political facts of a day; it also links global phenomena, such as climate change, to the day-to-day existence of people by reporting on their impact on individuals or communities. This strategy has been shown to capture and retain readership, and to enhance political engagement of the population (Cho and Gower, 2006; Zhao et al., 2019; Figenschou et al., 2023).

This “human face” of news reporting has been studied in the social sciences, ever since Semetko and Valkenburg (2000) defined the term **human interest (“HI”) framing** in their seminal paper as bringing “**a human face or an emotional angle to the presentation of an event, issue, or problem.**” News articles containing HI frames present an event, issue, or problem with an emphasis on the personal, emotional, and humanising aspects of the story. While NLP has formalized and quantified media framing from a variety of perspectives (Ali

¹HI labels, code to reproduce our data set and fine-tuned classifiers will be released upon acceptance.

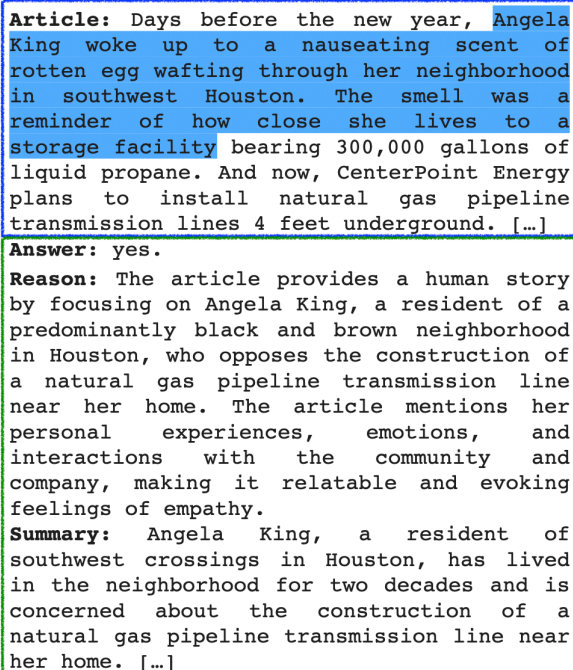


Figure 1: Top: A news article with Human Interest framing highlighted in blue (abbreviated for space reasons). Bottom: Output from our automated labelling system.

and Hassan, 2022; Vallejo et al., 2024), approaches have predominantly emphasised different framing devices and implications. We complement this line of work by presenting, to the best of our knowledge, the first large-scale, cross-cultural study of HI framing in news reporting.

Figure 1 shows an example of a news article that exhibits HI framing: the article about health and safety risks of a planned natural gas pipeline leads with a story about a local who is already suffering disruption from facilities in the area. The narration is vivid, referring to smells and health risks, and in doing so humanises the issue and evokes disgust and empathy in the reader, resulting in a connection with the affected community (Zhao et al., 2019).

In this paper, we study HI framing in the context of climate change reporting, analysing the specific

056 characteristics of HI framing (vs. non-HI articles) 106
057 in the context of this issue. Climate change, as a 107
058 pressing issue which effects individuals and com- 108
059 munities around the globe, is an omnipresent topic 109
060 in the news (Poushter et al., 2022) and a compelling 110
061 testbed for HI framing. Its nature as a global, slow- 111
062 moving issue with no clear single “face” renders 112
063 personalisation through individual impact stories 113
064 all the more important to engage and persuade the 114
065 audience of its imminent effects (Gibson et al., 115
066 2016). 116

067 We create a data set of climate change report- 117
068 ing from major news outlets in 2023, across four 118
069 English-speaking countries: Australia, India, the 119
070 USA, and the UK. We first label each article with 120
071 the presence or absence of a HI frame, using large 121
072 language models with human oversight. We then 122
073 devise a set of aspects to analyse the prevalence, 123
074 properties and cultural differences of HI framing, 124
075 including topics, victims and sentiment. Finally, 125
076 we supplement document embeddings from a fine- 126
077 tuned RoBERTa model (Liu et al., 2019b) with 127
078 these features to train a model that classifies HI 128
079 framing articles with high accuracy. 129

080 Our findings reveal global and culture-specific 130
081 characteristics of HI stories, with the USA having 131
082 the largest prominence of HI frames. HI frames 132
083 across all countries predominantly focus around 133
084 victims in the *general public*; are characterized by a 134
085 higher proportion of first- and third-person personal 135
086 pronouns; and exhibit more extreme (positive and 136
087 negative) sentiment compared to articles with no 137
088 HI frame. At the same time, the topic and victim 138
089 distribution in HI stories differs across countries. 139
090 Our framework and resources provides a basis for 140
091 future research into the manifestation and effects 141
092 of human interest framing. 142

093 2 Background 143

094 Our work spans HI framing, framing of climate 144
095 change and NLP models of framing. 145

096 **Human Interest Framing.** Previous studies have 146
097 illustrated positive and negative impacts of HI fram- 147
098 ing on the general public. Jebril et al. (2013) in- 148
099 vestigated the effect of *conflict* and *human interest* 149
100 frames on political knowledge gain. They found 150
101 that exposure to both frames increased political 151
102 knowledge, with the higher benefit for those with 152
103 low political interest, suggesting that HI framing 153
104 can be used as a strategy to educate about poli- 154
105 tics. Valenzuela et al. (2017) studied the effect of 155
156

106 HI framing on news sharing on Twitter and Face- 107
108 book, analysing news stories and interviewing dig- 108
109 ital journalists and editors. Although they report 109
110 low correlation of HI framing and news sharing, 110
111 their interviewees reported that such frames are 111
112 associated with highly viral stories and linked HI 112
113 framing to *emotional arousal*, suggesting that they 113
114 can evoke strong emotions and encourage sharing. 114
115 Another line of work found that HI framing may 115
116 amplify emotional responses and negatively im- 116
117 pact perceptions of crises, potentially amplifying 117
118 intensity of evaluations and perception of the is- 118
119 sue compared to other reporting styles (Cho and 119
120 Gower, 2006; Zhao et al., 2019). 120

120 **Framing climate change.** Research on climate 120
121 change framing has been extensive, with a strong 121
122 bias towards countries in the global north (She- 122
123 hata and Hopmann, 2012; Hart and Feldman, 2014; 123
124 O’Neill et al., 2015). Recent analyses of climate 124
125 change coverage across ten countries in the global 125
126 south and north (Hase et al., 2021) have found a 126
127 growing emphasis on societal aspects of climate 127
128 change, with the global south focusing on chal- 128
129 lenges and implications for society. In NLP, there 129
130 has been an increase in attention and research on 130
131 climate change in recent years, including releas- 131
132 ing datasets for climate claim verification (Diggel- 132
133 mann et al., 2020), exploring neutralisation (cli- 133
134 mate denial) techniques (Bhatia et al., 2021), ex- 134
135 tracting narratives (Stammach et al., 2022; Fr- 135
136 ermann et al., 2023), or automatically detecting net 136
137 zer or reduction targets (Schimanski et al., 2023). 137
138 Our research builds on the humanities studies by 138
139 devising a framework to expose HI frames at scale 139
140 with NLP methodology. It complements prior work 140
141 on climate framing in NLP with its focus on human 141
142 interest specifically. 142

143 **Framing in NLP** In NLP, a large number of 143
144 studies have investigated framing. Both super- 144
145 vised (Naderi and Hirst, 2017; Khanehzar et al., 145
146 2019; Huguet Cabot et al., 2020; Mendelsohn 146
147 et al., 2021) and unsupervised (Tsur et al., 2015; 147
148 Menini et al., 2017) approaches have been applied 148
149 to benchmarks including the Media Frames Cor- 149
150 pus (MFC) (Card et al., 2015) and Gun Violence 150
151 Frame Corpus (GVFC) (Liu et al., 2019a). How- 151
152 ever, Frermann et al. (2023) is the only study that 152
153 has previously investigated HI framing in NLP. 153
154 They released a corpus of 428 English news arti- 154
155 cles annotated with five frames (including HI), 155
156 finding that HI-stories are particularly prevalent in 156

right-leaning outlets. While we use some of their definitions and annotations for our analyses, our primary focus is on HI stories and their dynamics across countries and cultures. This study provides insights at a much larger scale with a corpus of more than 10k news articles about climate change from several countries. We also combine linguistic features derived from topic modelling, entity role analysis, and sentiment analysis with pre-trained models to improve model accuracy in identifying articles with HI framing. To the best of our knowledge, this is the first study in NLP investigating HI framing across several countries/cultures at scale.

3 Data Construction

In this section we describe the process of collecting and validating our cross-cultural data set of HI-stories. This process comprises three steps: (1) collecting a large number of news articles; (2) labelling the collection using *Mixtral-8x7B-Instruct-v0.1* (Jiang et al. (2024), hereafter referred to as Mixtral); and (3) manually validating the quality of the obtained labels. Below, we detail each of these steps.

3.1 Climate Change Corpus

We collected English news articles from 2023 via the ProQuest database² using *climate change* as a keyword. To ensure a representative sample, we retrieved data from the top outlets by circulation in each of four countries: Australia, India, UK, and USA.³ We obtained the political leaning for each of the outlets from the Media Bias Fact Check website⁴, ensuring that the outlets for each country span the political spectrum (full list of outlets in Appendix A, Table 3). After de-duplicating the articles by exact match, we removed documents with lengths outside the range of [100, 9000] words. This resulted in a total of 19,153 articles. Our strategy of matching articles by a single keyword (*climate change*) lead to a large number of articles that only tangentially concern the topic. We further filtered the articles by prompting Mixtral to return only those articles with climate change as the primary focus. We validated this step through manual inspection with a preference to err on the side of precision (i.e., remove relevant rather than retain ir-

²<https://www.proquest.com>

³Based on the Wikipedia list of top newspapers by circulation: <https://en.wikipedia.org/Newspapers>

⁴<https://mediabiasfactcheck.com>

Climate Change Corpus (CC-corpus)			
Country	Articles	HI-articles	% HI-articles
United States	3,836	1,078	28%
India	3,193	404	13%
Australia	2,007	318	16%
United Kingdom	1,264	270	21%
Total	10,300	2,070	20%
Control Corpus			
Country	Articles	HI-articles	% HI-articles
United States	3,762	1,517	40%
India	3,195	893	28%
Australia	1,993	706	35%
United Kingdom	1,267	373	29%
Total	10,217	3,489	33%

Table 1: Number of articles and proportion of HI-stories in the Climate-Change corpus (CC) (top) and control corpus (bottom).

relevant articles).⁵ Our final climate change corpus (“CC-corpus” hereafter) comprises 10,300 news articles. The distribution of articles across the four countries is shown in Table 1 (top, Articles).

Human Interest labels Given the article collection, we further prompt Mixtral to label articles as containing HI framing, or not. Our prompt included a definition of HI-stories based on Frermann et al. (2023) and Semetko and Valkenburg (2000), as well as one positive and two negative examples that were manually annotated by all authors. Mixtral is prompted to return: (1) a binary label indicating whether a given news article contains a HI-story, (2) an explanation for the label, and (3) a concise summary of the HI-story, if present. An article is considered to contain a HI-story when it presents a personal story of an individual, even if it is a short mention – we did not have length or prominence requirements.⁶ Figure 1 presents a news article with HI-story and the full Mixtral output.

3.2 Control Corpus

To gauge the extent to which our findings are specific to the issue of climate change, we collected a control sample of news articles from the same outlets and time period, but without any content restriction.⁷ The sample was down-sampled to match the size of the climate change dataset, ensuring uniform coverage over 2023 (see Table 1, bottom). This data was also labelled using Mixtral with the

⁵See Figure 17 in the Appendix for the full prompt

⁶See Figure 16 in Appendix Figure 16 for the full prompt.

⁷We did not provide any keyword in ProQuest.

same prompt as the CC-corpus (see Figure 16), resulting in 10,217 articles.

3.3 Human Annotation

To verify the HI labels predicted by Mixtral, we designed an annotation task using POTATO (Pei et al., 2022). Annotators first indicated whether the article contains a HI-story, using the same definition of HI-story used to prompt Mixtral, *without* seeing Mixtral’s prediction. Next, Mixtral’s original answer and reasoning were displayed and annotators were asked to indicate if they agree with this reasoning. Importantly, annotators were not able to change their initial assessment at this point. Finally, if present, Mixtral’s summary of the HI-story is presented, and annotators were again asked whether they consider that summary to be correct.

We assess the accuracy of our HI labels by manually verifying 100 Mixtral predictions. We recruited 10 local volunteer annotators, all with a graduate-level education, each of whom annotated 10 articles. Our human annotators agreed with Mixtral 81% of the time. In 12% of the cases, Mixtral classified an article as HI-story but annotators did not, and in the remaining 5% annotators categorised an article as HI but Mixtral did not. We consider this to be a strong result, given the complexity and ambiguity of the task.

4 Methodology

Recall that HI stories provide “a human face or an emotional angle to the presentation of an event, issue, or problem” (Semetko and Valkenburg, 2000). Departing from this definition, we define four measurable key aspects along which we analyze HI stories across cultures. In Section 5 we will show that these aspects are also empirically valuable for predicting the presence of HI stories in news articles. First, we analyse the *events, issues, or problems* by inspecting topics that dominate HI stories and discriminate between articles with and without HI stories. Second, given the *human* angle of HI stories we inspect the distribution of personal pronouns. Based on the same motivation, we thirdly analyse the types of victims portrayed in HI stories. Fourth, we assess the level of *emotion* in HI stories by comparing their sentiment intensity against non-HI articles.

4.1 Prevalence of HI framing

We analysed the prevalence of HI framing in climate change news reporting, compared to general

news. Out of the 10,300 climate change news articles, 2,070 were predicted by Mixtral to contain HI-stories (see Table 1, top), where we also show the number of HI-stories per country and their proportion. Considering the total number of HI framed articles in the CC-corpus (Table 1, top) and the Control-corpus (Table 1, bottom), we observe that HI framing is overall less prevalent in climate change reporting compared to the general news. This observation holds for every country in our study.

Comparing HI framing prevalence across countries, we observe that articles from the USA are more likely to contain HI framing than the other countries, but that the relative drop in prevalence over the control corpus is much higher than the UK. By contrast, India has the lowest proportion of HI-stories, and also the largest relative drop in HI-stories over the control corpus for that country, followed by Australia (with the second highest drop). These numbers demonstrate that there is a substantial difference in the prevalence of HI-stories in climate change news compared to general news in Australia and India, with a much lower proportion of HI-stories in climate change news.

4.2 Themes across countries and framings

How similar are the themes of climate change news reporting across the UK, USA, Australia, and India? And, how do those topics differ between articles with and without HI framing? To shed light on these questions, we conducted an exploratory analysis of the collected news articles with the help of a topic model based on BERTopic (Grootendorst, 2022). We applied this model on the CC articles and compared the obtained topics across countries and story-types.

First, we inspect the pairwise similarity of countries’ climate change reporting on the basis of their induced topic distribution. Figure 2 depicts the rank-biased overlap (RBO) (Webber et al., 2010) comparison across countries for articles *with* HI-stories.⁸ From the graph we observe that India is most dissimilar compared to the other three countries, while the United States, Australia, and the UK share a higher rank similarity. We provide a full list of representative words and topic ids in Appendix C. This finding suggests that India’s climate change reporting about *wild life* (topic 1), *Buddha*

⁸We also include the comparison for only the top 15 most common topics’ rank in Appendix C, with overall consistent results.



Figure 2: Heatmap over all topics’ rank similarity across countries. The image shows that India is less similar to all other countries, while the United States shares a high rank similarity with Australia and the United Kingdom.

(topic 68), and *agriculture* (topic 74) is not closely aligned with that of other countries, which overlap in topics such as *nature and environment* (topic 11), *electric cars* (topic 2) and *sea life* (topic 7). This could be due to various factors such as differences in climate change policies, cultural and religious influences that are reflected in the way news are reported, or the way climate change affects each of the countries. We provide the full list of representative words in HI-stories for the top 15 topics in each country (see Tables 4 to 7). We also applied the same analysis to the control corpus and obtained consistent results, with India again emerging as the most dissimilar country. It is noteworthy, however, that the control corpus revealed a higher overlap between Australia and the UK compared to the USA. This suggests that Australia and the UK have closely aligned agendas in general, but that they diverge in the specific domain of climate change.

Next, we turn to the difference in themes between HI and non-HI articles, pooling articles across all four countries. To this end, we fit a logistic regression model that predicts whether an article contains HI framing, with the topic distribution as input features.

Figure 3 depicts the most predictive features for the positive label (HI-framing; top) and the negative label (non-HI framing; bottom) in climate change stories, based on labels from Mixtral (see Table 8 for full list). We observe that the most common words in each topic without HI-stories are associated with corporate and financial issues,

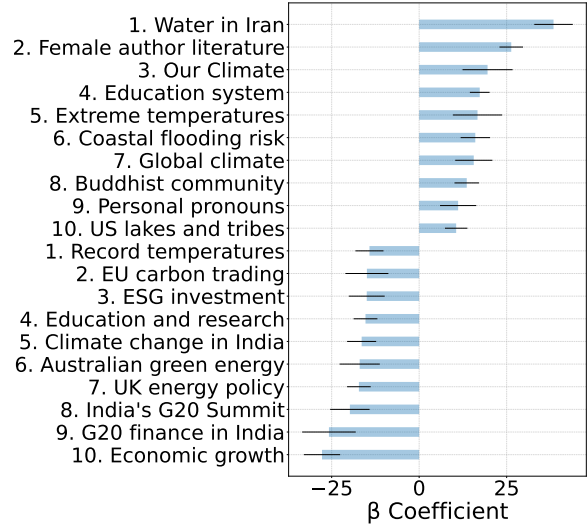


Figure 3: Logistic regression coefficients topic distributions in predicting articles containing HI-stories (positive/right) vs non HI-stories (negative/left).

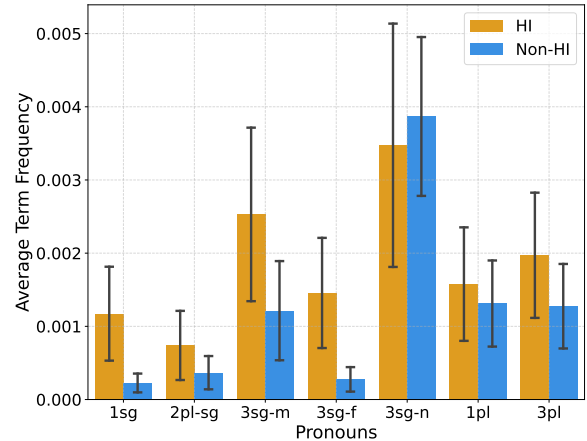


Figure 4: Comparison of averaged pronoun term frequency (including standard error of the mean) across HI- and non HI-stories. 1pl denotes first person plural, 1sg first person singular, etc.

international relations, and energy policy, whereas HI-stories are associated with environmental crises, education, personal opinions, and cultural/religious communities. Interestingly, several of the most predictive topics for HI-stories are dominated by personal pronouns, like in *our climate*, *female literature* and *personal pronouns* (see full top words in Table 8). We next inspect the role of pronouns in HI-stories in more detail.

4.3 Personal pronouns in HI-stories

To analyse pronoun usage, we categorised personal pronouns by person and number (1sg={*I, me, mine, my*}, 1pl={*we, us, our, ours*}, etc; full lists of pronouns per group are in Table 9). We then calculated

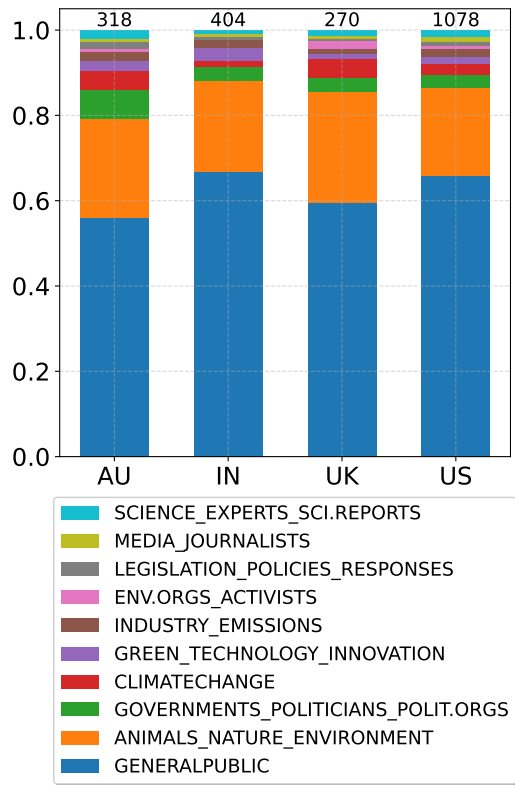


Figure 5: Distribution of entity groups portrayed as victims across countries.

the document term frequency for each group and compared the results across HI- and non HI-stories.

Figure 4 shows the results averaged per document, and reveal that pronouns are overall more prevalent in HI-stories (with the exception of the third-person neutral pronoun *it*). In contrast, third-person masculine and feminine and first-person singular pronouns are highly prominent in HI-stories which involve first-person narratives as quotes, and descriptions of individuals with corresponding third-person pronouns. Our analysis of the control corpus (see Figure 10) shows a consistent pattern, suggesting that pronouns are determinant in shaping the narrative of HI-stories, and that this observation extends beyond climate change.

4.4 Entity Analysis

The drama triangle theory, first proposed by Karpman (1968), describes the three primary roles that individuals can play in a conflict: the perpetrator (or villain); the victim, who is affected by the villain’s actions; and the rescuer, who acts as the victim’s helper. This theory was later adapted to analyse narrative roles in news stories (Gerbner and Gross, 1976; Van Dijk, 2013; Jones et al., 2014). In NLP, researchers have also explored

the automatic prediction of narrative roles (Jahan et al., 2021; Stambach et al., 2022), including in the context of climate change (Frermann et al., 2023). Here, we investigate the victim role in HI-stories about climate change, arguing that it is a crucial and emotionally evocative aspect of climate change narratives. We use the narrative frames corpus (Frermann et al., 2023), which includes entity group annotations (stakeholders in their paper), to analyse the character roles in these stories. The corpus categorises entities into ten entity groups, e.g. mapping *the earth* or *the oceans* to ANIMALS_NATURE_ENVIRONMENT, and *climate protests* to ENV.ORG_ACTIVIST. The original mapping was done by a human annotator.

Identifying these entity groups in each story enables us to find trends and patterns of entities being pictured as the victim in climate change articles across cultures. We automatically map victims in our HI-stories to their entity group with Mixtral, and use the gold mappings from Frermann et al. (2023) to validate our method. We prompted Mixtral to identify the most prominent entity portrayed as a victim in each HI-story, and mapped it to a climate change entity group from the narrative frames corpus.

Figure 5 illustrates the proportion of entity groups presented as victims in HI-stories across countries. Unsurprisingly, the GENERAL PUBLIC is the most prominent victim group, with higher proportions in India and the United States. The United Kingdom has the highest proportion of stories featuring ANIMALS AND NATURE as victims while Australia has the most articles depicting GOVERNMENT entities and SCIENCE EXPERTS as victims. This may be influenced by the coverage of Jacinda Ardern’s resignation as New Zealand’s Prime Minister, which was attributed to speculation about her *lack of delivery* on issues like climate, and the Labour party’s efforts to address issues like climate change to face election and avoid *suffering* the consequences. We also present this analysis across news outlets (see Figure 11 in Appendix E).

4.5 Sentiment Analysis

As quantified in previous studies from the communication and political communication, HI framing evokes emotions and is more prone to impact the behaviour of readers, such as article sharing in social media (Trifiro et al., 2023). Here, we investigate the difference of sentiment distributions over the averaged sentence polarity of individual articles.

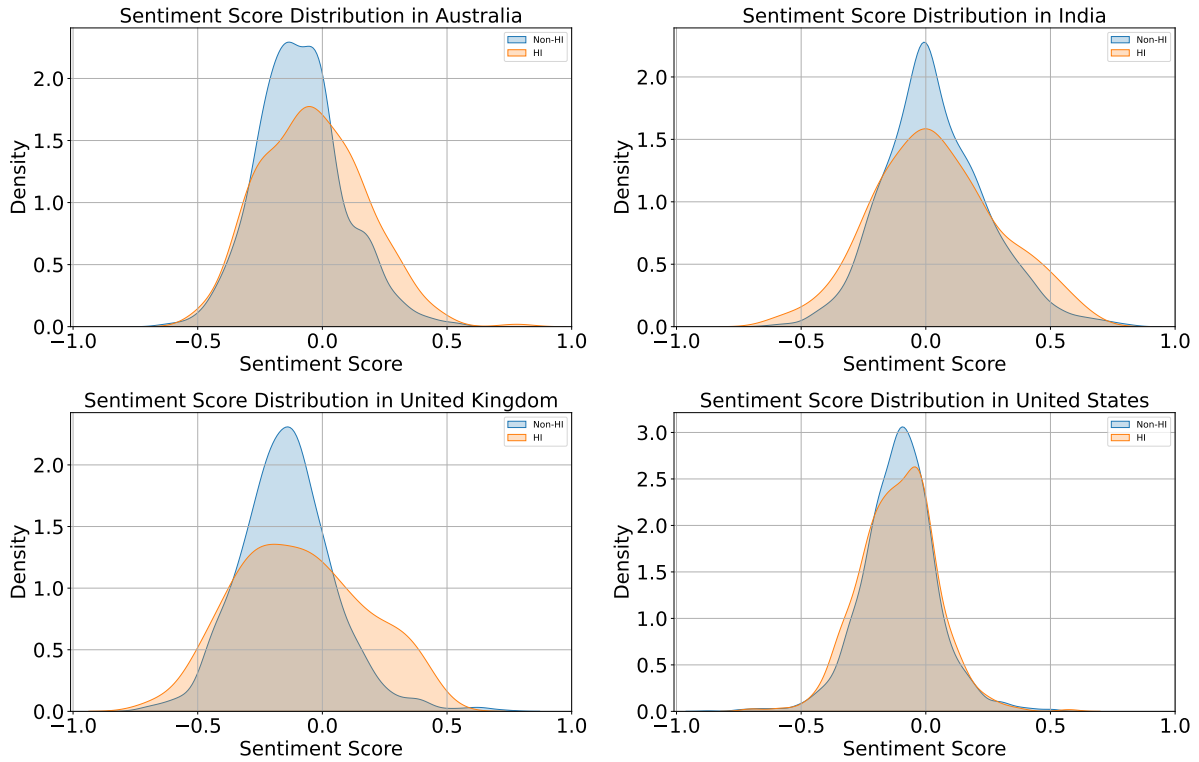


Figure 6: Comparison of sentiment analysis across countries. The x -axis denotes the average sentiment score, ranging from -1 (negative sentiment) to 1 (positive sentiment). A score of 0 indicates neutral sentiment.

We used a pre-trained version of RoBERTa (Liu et al., 2019b) —fine-tuned on Twitter data for sentence sentiment analysis (Barbieri et al., 2020) — to predict a sentiment score at the sentence level. These scores are averaged to obtain a score per article. Figure 6 presents a distribution comparison of HI and non-HI articles per country. We observe that in general, news articles tend to have a negative averaged sentiment for both categories, with the exception of India. This trend is stronger for non-HI-stories. This analysis also shows that HI-stories tend to have a wider range of sentiment, with Australia and the United Kingdom being slightly more biased towards positive sentiment. The fact that HI articles have a broader sentiment spectrum aligns with findings in previous research (Trifiro et al., 2023; Valenzuela et al., 2017) that this type of story tends to evoke emotion with the purpose of impacting readers’ behaviour. The control corpus exhibits similar patterns (see Figure 12), with minor differences in the breadth of sentiment in HI-stories for India and the UK.

5 HI Framing Classification

To evaluate the effectiveness of the analysed features, we conducted a final experiment to automat-

ically classify HI framing articles. We trained a single-layer multilayer perceptron (MLP) classifier using a combination of RoBERTa embeddings (Liu et al., 2019b) and linguistic features (topics, pronouns, and sentiment). We further fine-tuned a RoBERTa base (Liu et al., 2019b) model on 3,600 automatically-labelled articles, and use the resulting embeddings in a similar MLP model, combined with linguistic features. All models are tested on the human-labeled testset of 100 articles. The results, presented in Table 2 show, first, that it is possible to predict HI-stories in the news with a very high accuracy exceeding 90%. We further find that adding topic information improves the performance over direct RoBERTa fine-tuning,⁹ while the other features are less effective. The fine-tuned models outperform Mixtral.

6 Discussion

Our study reveals systematic differences in the way climate change is reported in news articles across the four investigated countries. The topic modelling results indicate that India stands out as the country with the lowest rank similarity across all

⁹The model will be release and we hope it will support further work on HI frames.

Model/Features	Accuracy	
Majority class	0.52	
<i>Mixtral-8x7B-Instruct-v0.1</i>	0.81	
MLP using:	Fine-Tuned	Off-the-Shelf
Embs	0.89	0.85
Embs + Topics	0.92	0.78
Embs + Sentiment	0.87	0.80
Embs + Sent. + Topics	0.89	0.77
Embs + Pronouns	0.86	0.83
Embs + Pron. + Topics	0.88	0.81

Table 2: Accuracy in classifying articles with and without HI framing on manual annotated data. Top: baseline, and *Mixtral-8x7B-Instruct-v0.1* labelling performance. Bottom: MLP trained with only embeddings (Embs.) and additional features. Embeddings were either extracted from a Fine-Tuned or an Off-the-Shelf version of RoBERTa.

four nations, which may reflect its distinct environmental context and socio-economic characteristics — as a global south country — compared to all the other global north countries (Billett, 2010). It is interesting to note that Australia and the UK are closer to the USA than to each other in terms of topic similarity, despite both nations sharing a common monarchy and historically being culturally closer. This result may be influenced by the fact that Australia and the USA are likely to be similarly impacted by climate change, with both countries experiencing significant environmental challenges in terms of adverse weather events and impacts on agriculture.

Our analysis of entities portrayed as victims in news articles exposes that the general public is the most frequently depicted group in a victim role across all four countries. This finding is consistent with previous research from the communication sciences, reporting the use of HI framing to humanise stories and influence public opinion (Zhao et al., 2019). It is noteworthy that Australia distinguishes itself from other countries by having the highest proportion of articles portraying governmental entities, legislation policies, and science experts as victims, which seems counter-intuitive given their typically authoritative roles. This observation leads to questions about the media’s portrayal of these groups in the context of climate change, and its potential impact on public perception of climate change.

Our sentiment analysis showed that non-HI climate change news articles tend to exhibit negative sentiment. This phenomenon is more evident in

Australia, UK, and USA, where sentiment analysis suggests a more pronounced negative tone in news articles. In contrast, HI-stories have a wider sentiment spectrum and lean towards more positive sentiment, specially in Australia and UK.

These findings lead us to better understand the use of the HI frame across cultures and expose patterns that can be useful for communication scientists and the public in general. From previous research we know that the way climate change is framed and reported can have a significant impact on public perceptions and attitudes towards the issue (Cho and Gower, 2006). By understanding differences, as shown in our study, we can propose and develop tools to mitigate possible reporting biases.

7 Conclusion

Our cross-cultural analysis of climate change HI-stories in news articles has provided insights into the ways in which climate change is framed and reported in different parts of the world. Through topic modelling, entity, and sentiment analyses we were able to expose systematic differences in news articles across the four nations. Our study revealed interesting findings like the diversity of topics across countries, the different victim roles in all countries, and the broader sentiment in HI compared to non HI-stories.

Overall, our study demonstrates the value of combining multiple NLP approaches to gain a deeper understanding of complex issues like climate change. By integrating topic modelling, entity analysis, and sentiment analysis, we can gain a more comprehensive picture of how climate change is framed and reported in different parts of the world.

8 Limitations

The scope of this work is limited to news articles in English from the listed outlets in Australia, India, UK, and USA. We are aware that our work does not necessarily represent each country’s perspective but rather those of the examined outlets. We are also aware that large language models manifest a wide variety of biases that we might not be aware of, and that could be harmful. We hope that with this research we can shed light on systematic differences in the use of HI framing in different countries in English. We also understand that ProQuest data is licensed and that reproducibility might be af-

583	fected. We will release ProQuest ids along with		
584	labels and generated data for further research		
585	References		
586	Mohammad Ali and Naemul Hassan. 2022. A survey of computational framing analysis approaches .		
587	In <i>Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing</i> , pages		
588	9335–9348, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.		
589			
590			
591			
592	Francesco Barbieri, Jose Camacho-Collados, Luis Espinosa Anke, and Leonardo Neves. 2020. TweetEval: Unified benchmark and comparative evaluation for tweet classification .		
593	In <i>Findings of the Association for Computational Linguistics: EMNLP 2020</i> , pages		
594	1644–1650, Online. Association for Computational Linguistics.		
595			
596			
597			
598			
599	Shraey Bhatia, Jey Han Lau, and Timothy Baldwin. 2021. Automatic classification of neutralization techniques in the narrative of climate change scepticism .		
600	In <i>Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies</i> , pages		
601	2167–2175, Online. Association for Computational Linguistics.		
602			
603			
604			
605			
606			
607	Simon Billett. 2010. Dividing climate change: global warming in the indian mass media . <i>Climatic Change</i> ,		
608	99:1–16.		
609			
610	Dallas Card, Amber E. Boydston, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2015. The media frames corpus: Annotations of frames across issues . In <i>Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)</i> , pages		
611	438–444, Beijing, China. Association for Computational Linguistics.		
612			
613			
614			
615			
616			
617			
618			
619	Seung Ho Cho and Karla K. Gower. 2006. Framing effect on the public’s response to crisis: Human interest frame and crisis type influencing responsibility and blame . <i>Public Relations Review</i> , 32(4):420–422.		
620			
621			
622			
623	Thomas Diggelmann, Jordan L. Boyd-Graber, Jannis Bulian, Massimiliano Ciaramita, and Markus Leipold. 2020. CLIMATE-FEVER: A dataset for verification of real-world climate claims . In <i>Proceedings of the Tackling Climate Change with Machine Learning workshop at NeurIPS</i> , Online.		
624			
625			
626			
627			
628			
629	Tine Ustad Figenschou, Kjersti Thorbjørnsrud, and Daniel C Hallin. 2023. Whose stories are told and who is made responsible? human-interest framing in health journalism in norway, spain, the u.k. and the u.s. <i>Journalism</i> , 24(1):3–21.		
630			
631			
632			
633			
634	Lea Frermann, Jiatong Li, Shima Khanehazar, and Gosia Mikolajczak. 2023. Conflicts, villains, resolutions: Towards models of narrative media framing . In <i>Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)</i> , pages	636	
	8712–8732, Toronto, Canada. Association for Computational Linguistics.	637	
		638	
		639	
		640	
	George Gerbner and Larry Gross. 1976. The scary world of tv’s heavy viewer. <i>Psychology today</i> , 9(11):41–45.	641	
		642	
		643	
	Timothy A Gibson, Richard T Craig, Allison C Harper, and Jordan M Alpert. 2016. Covering global warming in dubious times: Environmental reporters in the new media ecosystem. <i>Journalism</i> , 17(4):417–434.	644	
		645	
		646	
		647	
	Maarten Grootendorst. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure . <i>CoRR</i> , abs/2203.05794.	648	
		649	
		650	
	P. Sol Hart and Lauren Feldman. 2014. Threat without efficacy? climate change on u.s. network news . <i>Science Communication</i> , 36(3):325–351.	651	
		652	
		653	
	Valerie Hase, Daniela Mahl, Mike S. Schäfer, and Tobias R. Keller. 2021. Climate change in news media across the globe: An automated analysis of issue attention and themes in climate change coverage in 10 countries (2006–2018) . <i>Global Environmental Change</i> , 70:102353.	654	
		655	
		656	
		657	
		658	
		659	
	Pere-Lluís Huguet Cabot, Verna Dankers, David Abadi, Agneta Fischer, and Ekaterina Shutova. 2020. The Pragmatics behind Politics: Modelling Metaphor, Framing and Emotion in Political Discourse . In <i>Findings of the Association for Computational Linguistics: EMNLP 2020</i> , pages	660	
	4479–4488, Online. Association for Computational Linguistics.	661	
		662	
		663	
		664	
		665	
		666	
	Labiba Jahan, Rahul Mittal, and Mark Finlayson. 2021. Inducing stereotypical character roles from plot structure . In <i>Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing</i> , pages	667	
	492–497, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.	668	
		669	
		670	
		671	
		672	
	Nael Jebri, Claes H. de Vreese, Arjen van Dalen, and Erik Albæk. 2013. The effects of human interest and conflict news frames on the dynamics of political knowledge gains: Evidence from a cross-national study . <i>Scandinavian Political Studies</i> , 36(3):201–226.	673	
		674	
		675	
		676	
		677	
		678	
	Albert Q. Jiang, Alexandre Sablayrolles, Antoine Roux, Arthur Mensch, Blanche Savary, Chris Bamford, Devendra Singh Chaplot, Diego de Las Casas, Emma Bou Hanna, Florian Bressand, Gianna Lengyel, Guillaume Bour, Guillaume Lample, Léo Renard Lavaud, Lucile Saulnier, Marie-Anne Lachaux, Pierre Stock, Sandeep Subramanian, Sophia Yang, Szymon Antoniak, Teven Le Scao, Théophile Gervet, Thibaut Lavril, Thomas Wang, Timothée Lacroix, and William El Sayed. 2024. Mixtral of experts . <i>CoRR</i> , abs/2401.04088.	679	
		680	
		681	
		682	
		683	
		684	
		685	
		686	
		687	
		688	
		689	
634	Michael D. Jones, Mark K. McBeth, and Elizabeth A. Shanahan. 2014. Introducing the Narrative Policy	690	
635		691	

692	<i>Framework</i> , page 1–25. Palgrave Macmillan US, New York.	
693		
694	Stephen Karpman. 1968. Fairy tales and script drama analysis. <i>Transactional analysis bulletin</i> , 7(26):39–43.	
695		
696		
697	Shima Khanehzar, Andrew Turpin, and Gosia Mikolajczak. 2019. Modeling political framing across policy issues and contexts . In <i>Proceedings of the 17th Annual Workshop of the Australasian Language Technology Association</i> , pages 61–66, Sydney, Australia. Australasian Language Technology Association.	
698		
699		
700		
701		
702		
703	Siyi Liu, Lei Guo, Kate Mays, Margrit Betke, and Derry Tanti Wijaya. 2019a. Detecting frames in news headlines and its application to analyzing news framing trends surrounding U.S. gun violence . In <i>Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)</i> , pages 504–514, Hong Kong, China. Association for Computational Linguistics.	
704		
705		
706		
707		
708		
709		
710		
711	Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019b. Roberta: A robustly optimized BERT pretraining approach . <i>CoRR</i> , abs/1907.11692.	
712		
713		
714		
715		
716	Julia Mendelsohn, Ceren Budak, and David Jurgens. 2021. Modeling framing in immigration discourse on social media . In <i>Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies</i> , pages 2219–2263, Online. Association for Computational Linguistics.	
717		
718		
719		
720		
721		
722		
723	Stefano Menini, Federico Nanni, Simone Paolo Ponzetto, and Sara Tonelli. 2017. Topic-based agreement and disagreement in US electoral manifestos . In <i>Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing</i> , pages 2938–2944, Copenhagen, Denmark. Association for Computational Linguistics.	
724		
725		
726		
727		
728		
729		
730	Nona Naderi and Graeme Hirst. 2017. Classifying frames at the sentence level in news articles . In <i>Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017</i> , pages 536–542, Varna, Bulgaria. INCOMA Ltd.	
731		
732		
733		
734		
735		
736	Saffron J. O’Neill, Hywel T. P. Williams, Tim Kurz, Bouke Wiersma, and Maxwell Boykoff. 2015. Dominant frames in legacy and social media coverage of the ipcc fifth assessment report . <i>Nature Climate Change</i> , 5:380–385.	
737		
738		
739		
740		
741	Jiaxin Pei, Aparna Ananthasubramaniam, Xingyao Wang, Naitian Zhou, Apostolos Dedeloudis, Jackson Sargent, and David Jurgens. 2022. POTATO: The portable text annotation tool . In <i>Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations</i> , pages 327–337, Abu Dhabi, UAE. Association for Computational Linguistics.	
742		
743		
744		
745		
746		
747		
748		
	Jacob Poushter, Moira Fagan, and Sneha Gubbala. 2022. Climate change remains top global threat across 19-country survey. <i>Pew Research Center</i> .	749 750 751
	Tobias Schimanski, Julia Bingler, Mathias Kraus, Camilla Hyslop, and Markus Leippold. 2023. ClimateBERT-NetZero: Detecting and assessing net zero and reduction targets . In <i>Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing</i> , pages 15745–15756, Singapore. Association for Computational Linguistics.	752 753 754 755 756 757 758
	Holli A. Semetko and Patti M. Valkenburg. 2000. Framing european politics: a content analysis of press and television news . <i>Journal of Communication</i> , 50:93–109.	759 760 761 762
	Adam Shehata and David Nicolas Hopmann. 2012. Framing climate change . <i>Journalism Studies</i> , 13:175–192.	763 764 765
	Dominik Stambach, Maria Antoniak, and Elliott Ash. 2022. Heroes, villains, and victims, and GPT-3: Automated extraction of character roles without training data . In <i>Proceedings of the 4th Workshop of Narrative Understanding (WNU2022)</i> , pages 47–56, Seattle, United States. Association for Computational Linguistics.	766 767 768 769 770 771 772
	Briana M. Trifiro, Akanksa Upadachay, Siyu Liu, Zhimin Zhang, and Kelsey Prena. 2023. Framing facebook content: An analysis of how framing strategies drive facebook engagement . <i>Journal of Promotion Management</i> , 29(8):1229–1249.	773 774 775 776 777
	Oren Tsur, Dan Calacci, and David Lazer. 2015. A frame of mind: Using statistical models for detection of framing and agenda setting campaigns . In <i>Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)</i> , pages 1629–1638, Beijing, China. Association for Computational Linguistics.	778 779 780 781 782 783 784 785 786
	Sebastián Valenzuela, Martina Piña, and Josefina Ramírez. 2017. Behavioral effects of framing on social media users: How conflict, economic, human interest, and morality frames drive news sharing . <i>Journal of Communication</i> , 67(5):803–826.	787 788 789 790 791
	Gisela Vallejo, Timothy Baldwin, and Lea Frermann. 2024. Connecting the dots in news analysis: Bridging the cross-disciplinary disparities in media bias and framing . In <i>Proceedings of the Sixth Workshop on Natural Language Processing and Computational Social Science (NLP+CSS 2024)</i> , pages 16–31, Mexico City, Mexico. Association for Computational Linguistics.	792 793 794 795 796 797 798 799
	Teun A. Van Dijk. 2013. <i>News Analysis</i> , 1st edition. Routledge.	800 801
	William Webber, Alistair Moffat, and Justin Zobel. 2010. A similarity measure for indefinite rankings . <i>ACM Transactions on Information Systems (TOIS)</i> , 28(4):1–38.	802 803 804 805

Xinyan Zhao, Mengqi Monica Zhan, and Brooke Fisher Liu. 2019. [Understanding motivated publics during disasters: Examining message functions, frames, and styles of social media influentials and followers](#). *Journal of Contingencies and Crisis Management*, 27(4):387–399.

A Data Statistics

Our dataset comprises articles from four countries: Australia, India, the UK, and the USA. To ensure a representative sample, we sourced articles from top outlets by circulation, which resulted in a diverse collection with a range of political leanings, from right to centre-left, as illustrated in Table 3. It includes also total of articles in both the climate change corpus and the control corpus and the news outlets by country.

Country	Outlet	CC	BG	Bias
AU	The Australian	1008	1008	Right-Center
	Sydney Morning Herald	386	372	Left-Center
	The Age	359	359	Left-Center
	The Daily Telegraph	124	124	Right
	Herald Sun	130	130	Right-Center
IN	The Times of India	2033	2033	Right-Center
	The Economic Times	337	337	Right-Center
	The Hindustan Times	823	825	Left-Center
UK	Daily Mail	502	504	Right
	The Daily Telegraph	470	470	Right
	The Daily Mirror	292	293	Left-Center
US	New York Times	1552	1549	Left-Center
	The Washington Post	709	710	Left-Center
	Wall Street Journal	791	791	Right-Center
	USA Today	572	572	Left-Center
	Los Angeles Times	212	140	Left-Center
	Total		10300	10217

Table 3: Number of articles across outlets including the Climate-Change corpus (CC) and our background corpus (BG) for comparison. The Bias column represents the bias rating extracted from Media Bias Fact Check.

B Experimental Setup Details

All our automatic annotations were done by prompting *Mixtral-8x7B-Instruct-v0.1* (Jiang et al. (2024)) using HuggingFace’s Serverless Inference API ¹⁰. We fine-tuned RoBERTa (Liu et al., 2019b) using a GPU *NVIDIA GeForce GTX 1080* for less than an hour. Our fine-tuning data comprises 3, 600 samples divided into 2144 articles for training and twice 528 for development and training. We used a batch size of 4 articles, truncated articles to 512 tokens and fine-tuned for 3 epochs. Our MLP clas-

¹⁰<https://huggingface.co/docs/api-inference/index>

sifier was implemented using *scikit-learn* ¹¹ and we trained for 10 to 20 epochs using early stopping.

C Topic Modelling Details

We conducted a cross-country comparison of the rank-biased overlap (RBO) in the top 15 most common topics (Figure 7) and the control corpus (Figure 8). Interestingly, the top 15 topics yielded consistent results with the broader analysis. However, results on the control corpus revealed a higher degree of overlap between the UK and Australia, when compared to the USA, recall that this is opposed to the results with climate change topics (Figure 2). This finding suggests that the UK and Australia share a high degree of similarity in their agendas, although this convergence is not the case in the domain of climate change.



Figure 7: Heatmap of Rank Similarity for the top 15 topics across countries.

We list top 15 topics in each country’s HI articles (Tables 4 to 7). These lists depict some overlap, specially between Australia (Table 4) and UK (Table 6), and UK (Table 6) and USA (Table 7).

In Table 8, we present the list of words that best capture the most predictive features for classifying articles HI framing using logistic regression. These words were used to automatically generate (using Mixtral) the topic names displayed in Figure 3. For comparison purposes we replicated the experiment using the control corpus to fit a logistic regression model. The results of are depicted in Figure 9. Here we observe that HI framing is related to *health*, *families* and *housing* while in

¹¹https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html



Figure 8: Heatmap of Rank Similarity across countries in the control corpus.

Top words in HI topics in Australia
94. our we climate is australia to and it of the
5. pacific australia albanese tuvalu islands defence minister china
50. students school education teachers university and they of
60. we our of it that and the about my is
14. energy power renewable coal electricity bowen grid nsw government
7. fish ocean coral whales marine reef salmon sea the fishing
6. greens mechanism labor government emissions bowen coal cent
70. australia energy government to and will zero investment the
75. museum art work she of music artist and with biennale
25. population productivity tax our government chalmers growth we
57: coastal homes flood the beach and to of erosion sea
11: trees forest tree forests and of the to in plants
43: wind offshore turbines farms project power farm energy orsted
34: king charles his queen royal he prince monarch palace speech
79: biodiversity land carbon farmers clearing cattle agriculture and to

Table 4: List of representative words in Australia’s HI articles for the top 15 topics (left: topic id).

Top words in HI topics in India
41. film films art delhi festival my is and we about
1. forest tiger conservation the of ministry wildlife project and
19. climate change environment and students state of india development
74. farmers agriculture farming millets rice food soil millet of
30. shimla himachal joshimath river the rs state hill in of
68. ladakh nuns buddha buddhist wangchuk drukpa nun sect mountains
99. metro transport cycling vehicles trams buses tram city will bus
4. india energy green hydrogen renewable and of solar the is
8. rainfall monsoon india imd normal temperature rain in nino recorded
38. education students university research india indian ip institute
114: youth y20 summit india modi lawson themes young and gen
36: mangrove mangroves fisheries coastal sand marine goa india and the
26: wheat prices crop crops farmers tea cotton production in food
87: earth day environment sustainable planet our waste awareness plastic
44: heat temperature heatwave india urban temperatures city and delhi

Table 5: List of representative words in India’s HI articles for the top 15 topics (left: topic id).

non HI-stories articles we see again *corporate* and *finance* topics. Interestingly, *security*, *police* and *violence* were not highly prominent for topics in articles related to climate change.

Top words in HI topics in UK
13. protesters police protest activists was stop he they the
0. labour sunak uk zero net oil to government gas the
11. trees forest tree forests and of the to in plants
34. king charles his queen royal he prince monarch speech
33. birds bees species bird they honey insects bee are cats
31. ireland irish we it to be is that the ryan
104. polar bears inuit arctic bear ice rangers liz they
95. we our of it that and the about my is
60. earthshot prince william awards winners singapore ceremony
28. weather temperatures uk summer we in record met more
7: fish ocean coral whales marine reef salmon sea the fishing
110: her she novel book you with it of and my
94: our we climate is australia to and it of the
119: heat europe temperatures sardinia italy in summer rome cool
2: electric vehicles cars car charging ev battery evs sales

Table 6: List of representative words in UK’s HI articles for the top 15 topics (left: topic id).

Top words in HI topics in USA
7: fish ocean coral whales marine reef salmon sea the fishing
16: heat temperatures degrees texas air hot conditioning in
22: iran water the in of iraq people and to dams
11: trees forest tree forests and of the to in plants
32: fire maui hawaii lahaina hawaiian was wildfire the and
12: food meat methane waste to of and farmers the dairy
48: flood flooding city storm rain york new water in the
60: we our of it that and the about my is
13: protesters police protest activists was stop he they the
35: trump taylor biden voters he his wilson ramaswamy you
40: fire canada wildfire smoke forest burned canadian of
2: electric vehicles cars car charging ev battery evs sales
18: california water drought storms rain county rivers snow
24: air smoke quality wildfire health canada york pollution the
52: montana court case plaintiffs that trial lawsuit the in

Table 7: List of representative words in USA’s HI articles for the top 15 topics (left: topic id).

Topics classified as HI articles
1. iran water the in of iraq people and to dams
2. her she novel book you with it of and my
3. our we climate is australia to and it of the
4. students school education teachers university and they
5. heat temperatures degrees texas air hot conditioning
6. coastal homes flood the beach and to of erosion sea
7. countries africa world climate developing billion bank
8. ladakh nuns buddha buddhist wangchuk nunnerly drukpa
9. we our of it that and the about my is
10. lake water river salt mississippi klamath duluth corps tribes
Topics classified as non-HI articles
1. record temperatures heat hottest july warmest degrees global
2. eu cbam trade wto carbon tax india border steel imports
3. esg companies funds investors financial investment corporate
4. education students university research india ip institute
5. climate change environment and students state of india
6. gas greens projects santos australian supply mechanism energy
7. labour sunak uk zero net oil to government gas the
8. india modi g20 summit ukraine g7 global russia pm issues
9. finance bank g20 india development mdb multilateral financial
10. population productivity tax our government chalmers growth we

Table 8: List of top words in each topic, based on these we automatically generated the names in Figure 3

D Pronoun Analysis Details

We present below the mapping list of pronouns across HI and non-HI stories (Table 9), as well as a comparison of pronoun prevalence in the control corpus (Figure 10). The results of this analysis

863
864
865
866

867
868
869
870
871

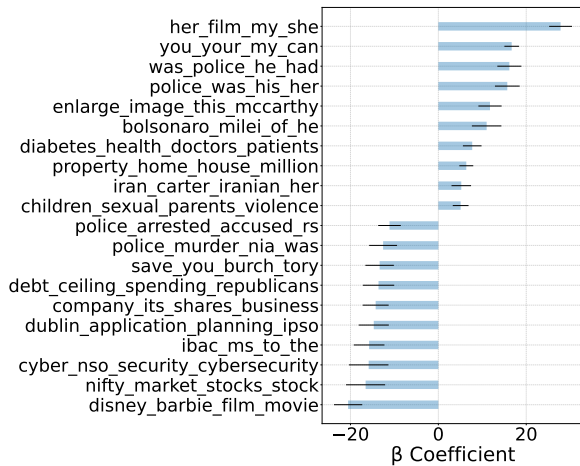


Figure 9: Logistic regression coefficients topic distribution in predicting HI framing articles using the control corpus.

Category Group	Pronouns
1sg	I, me, my, mine
2sg-pl.	you, your, yours
3sg-m	he, him, his
3sg-f	she, her, hers,
3sg-n	it, its
1pl	we, us, our, ours
3pl	they, them, their, theirs

Table 9: List of pronoun categories as represented in Figure 4.

align with our previous findings, indicating that the use of gendered third- and first-person pronouns is strongly associated with the presence of HI framing in an article.

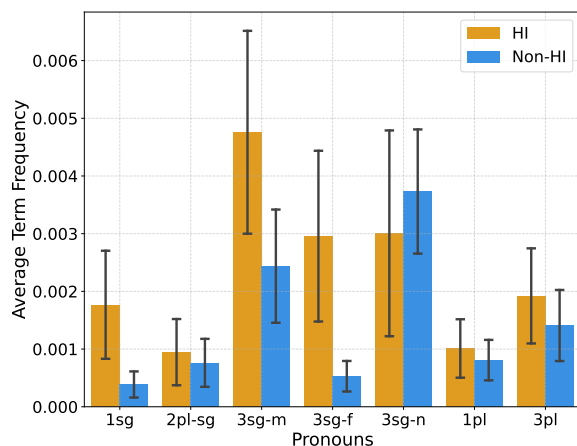


Figure 10: Pronoun term frequency compared across HI- and non-HI stories in the control corpus.

E Narrative Roles Details

We conducted our analysis of victim roles across news outlets too. The results, as depicted in Figure 11, show that certain entity groups are under-represented in specific outlets. This pattern may be attributed to the varying number of articles by each outlet.

F Sentiment Analysis Details

Figure 12 presents the distributions of sentiment analysis scores across countries for articles with and without HI framing. Results are consistent with the sentiment analysis conducted on the CC. We observed a slightly broader range of sentiment in all countries in HI-stories. Although the effect is more pronounced in climate change articles, this trend is also evident in the current analysis.

G Annotation Details

Our annotation task includes examples and reasons why articles should and should not be labelled as containing HI-stories. We also included a training phase in which we provide annotators with more examples. In this section, we provide a visual overview of the platform and task through a series of screenshots. Figure 13 shows the annotation introduction page, where annotators were presented with a clear explanation of the task requirements, including the number of samples to be annotated (10 articles and 3 training samples). All annotators agreed to the use of this data for research purpose.

Figure 14 provides examples of articles with and without HI framing, accompanied by explanations of the classification rationale and highlighting the key signals to consider. Figure 15 illustrates a sample annotation from the training phase, which mirrors the format of the actual examples. To minimise annotator fatigue, we deliberately selected shorter examples for the training phase.

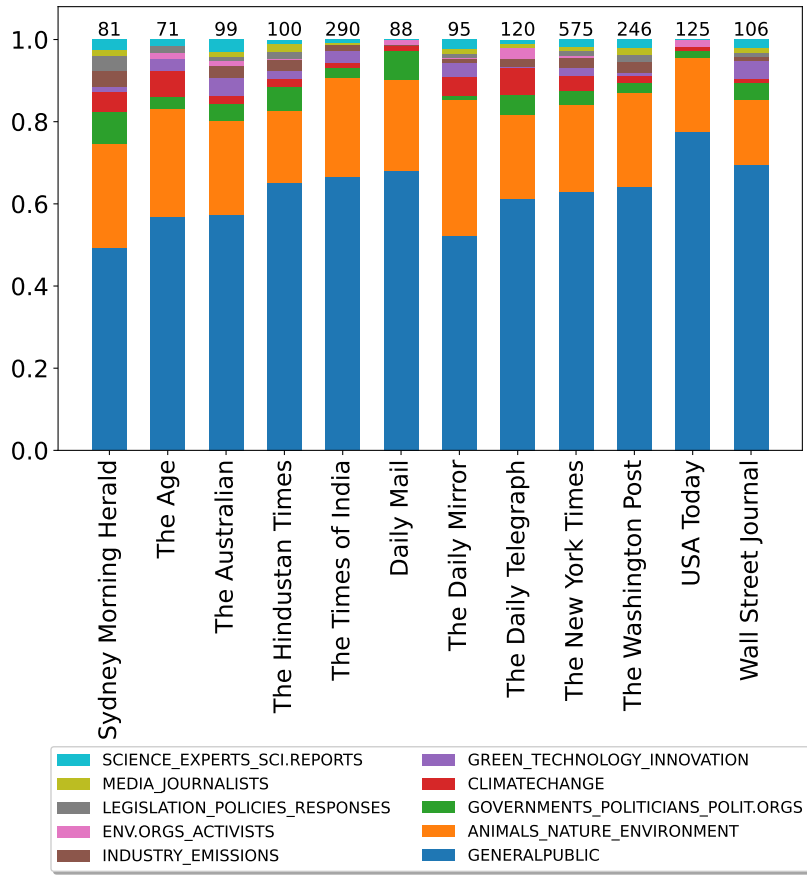


Figure 11: Distribution of entity groups portrayed as victims across outlets.

H Prompting Details

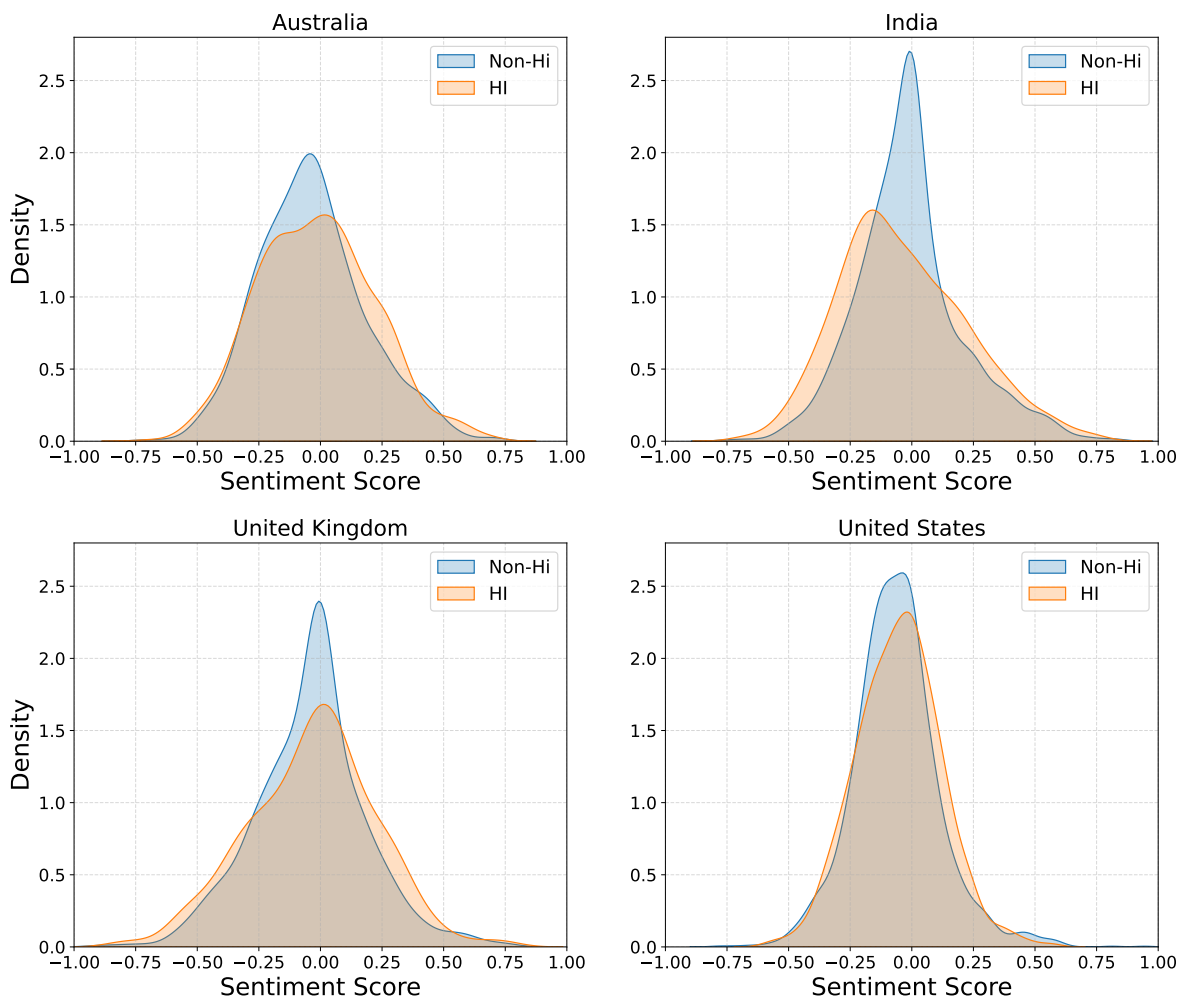


Figure 12: Comparison of sentiment score distribution in the control corpus across the four countries: Australia, India, the United Kingdom and the United States.

Finding Human Personal Stories
Home Statistics Help Finished 0
Currently logged in as

Plain Language Statement

Procedures:
 In this study, you will be presented with three examples to practice and ten news articles (real annotation) broadly related to climate change, one at a time. 13 articles in total, but only 10 real annotations. Your task is to determine whether the article contains a "Human personal story". After supplying your answer, you will be asked to review another assessment, including reasons and a summary of the personal story. You will then be asked whether you agree with their assessment or not. Please note that once you select your answer, you will not be able to change it.

Definition:
 "Human personal stories" are stories that contain private or personal details of the individuals involved. An article that contains a human personal story should reference the personal impact on one or more concrete, individual entities. See examples in the next page.

Risks:
 The risks and discomfort associated with participation in this study are no greater than those ordinarily encountered in daily life, such as when surfing the internet. There is a possibility of disagreement with the article's point of view, but please try to be as objective as possible in finding human personal stories.

Benefits:
 Your participation will contribute to a better understanding of how personal interest stories are used in news reporting on climate change. This understanding might help to expose media bias, which results in a more informed society.

Move forward

Figure 13: Introduction to the annotation task.

Examples

Example of an article with Human personal stories

- **Article:** Taylor Swift has sparked a wave of criticism online for the extraordinary carbon footprint she created using a private jet during the Australian leg of her Eras Tour. The megastar, 33, flew into Australia in her Bombardier Global 6000 earlier this month, and then used the luxury aircraft to travel between Melbourne and Sydney as she continued her tour. Swift also sent the aircraft to Hawaii to pick up her boyfriend Travis Kelce last week, transporting the NFL star to Sydney so he could watch her perform. On Tuesday morning, the Shake It Off hit-maker and her entourage jetted out of Australia using the same aircraft, en route to Singapore. Australian critics have since slammed Swift - who previously described climate change as a 'horrific situation' plaguing the world - for using a gas-guzzling jet rather than flying commercial.
 - **Reason:** The human personal story element comes from the mention of Taylor Swift's boyfriend, Travis Kelce, and how she sent her private jet to pick him up from Hawaii so he could watch her perform in Sydney. **Personal or private details include but not only, age, partner, address, details about the place of residence, financial information, sexual orientation, emotions about the individuals involved, etc.**
- **Article:** As a child, Abdallah Tabet would leave his home in Beirut during the summer to visit his grandmother in the lush mountains. When he'd return to the city, it felt like a 'concrete jungle' in comparison, he says, leaving him longing for the sight of trees. "It's funny how you appreciate things more when you don't have them," Tabet says. Several decades later, while he was living in Philadelphia and working as a landscape architect, that feeling resurfaced. For a while, Tabet had lived on the west side of the city, which was filled with tall sycamores, poplars and maples. But when he moved to the Graduate Hospital neighborhood in 2013, it seemed there were "almost no trees." Motivated to change that, Tabet volunteered to be a "tree tender" with the Pennsylvania Horticultural Society. Now he's the leader of his neighborhood's tree tender group, which gathers several times a year to plant and help maintain local trees. "It feels very rewarding," he says. "All the volunteers that I work with ... we're contributing to a big positive change for the city."
 - **Reason:** The story shares Abdallah Tabet's **personal experiences and feelings, such as his childhood memories of visiting his grandmother** in the mountains and his sense of longing for trees when he returned to the city. The text also provides specific details about Tabet's life, such as where he lived and worked, and how he became involved with the Pennsylvania Horticultural Society. **Relationships like Abdallah's to his grandmother are private relationships to the story, which points to a personal detail of the main person in the story.**

Figure 14: Examples presented in the annotation task.

Please read the following article carefully. Your task is to assess whether an article provides a "Human personal story". This means to examine if the article provides a human example or "human face" on the story's issue or whether the story goes into the private or personal lives of the entities involved. A strategy to find Human personal stories is to examine whether the story includes personal or private details like, but not only, age, partner, address, details about the place of residence, financial information, sexual orientation, emotions about the individuals involved, etc. After that you will be presented with an assessment and a summary done by someone else, please answer whether you agree with that assessment or not. Once you give an answer you won't be able to change your choice:

In a second step, you will be presented with an assessment and a summary done by another annotator, and you need to indicate whether you agree with that assessment or not. Optionally, you may comment on your answer.

Julián Castro, the first 2020 Democratic presidential candidate at the CNN climate town hall, the self-professed "climate candidate" whose six-part plan is said: "When I was mayor of San Antonio, I did believe that there were opportunities to be had with fracking that was going on in South Texas.". "We're coming to the end of the bridge," Mr. Castro said. He said while he isn't calling for an immediate national ban on fracking, he supports local communities that do so and called for moving to cleaner sources. Mr. Castro won some of his strongest applause with the need to protect low-income communities from environmental damages.

Recall that human personal stories include personal anecdotes, emotional language, or insights into individuals' private lives. Please select from the following options:

- Yes, there are personal stories in the article
- No, there are no personal stories are presented in the article

Figure 15: First question in annotation task.

[INST] You are a journalist analyst, and your task is to assess whether an article provides a human story. This means to examine if the article provides a human example or "human face" on the story's issue or whether the story goes into the private or personal lives of the entities involved. Your output should be ANSWER: yes, or ANSWER: no. Please provide your REASON justifying your answer. In the case of answering yes, give a short SUMMARY of the sentences containing the human example or private story. Here are three examples, one positive and two negatives. [/INST]

[EX] EXAMPLE1:

ARTICLE: Prince Harry defends use of private jets: "It's to keep my family safe" The Duke of Sussex has defended his use of private jets. "I came here by commercial. I spend 99% of my life travelling the world by commercial," he said. "Occasionally there needs to be an opportunity based on a unique circumstance to ensure that my family are safe - it's generally as simple as that." During his opening speech, the duke said "no one is perfect" in terms of his impact on the environment needs to ensure his family is safe.

ANSWER: yes

REASON: The article provides a human story by mentioning the Duke of Sussex's use of private jets to ensure his family's safety, triggering feelings of empathy and understanding.

SUMMARY: In the article, Prince Harry defends his use of private jets, stating that he primarily travels commercially but occasionally opts for private jets for safety reasons to protect his family. He acknowledges his environmental impact but emphasizes the priority of ensuring his family's safety.

EXAMPLE2:

ARTICLE: Julián Castro, the first 2020 Democratic presidential candidate at the CNN climate town hall, the self-professed "climate candidate" whose six-part plan is said: "When I was mayor of San Antonio, I did believe that there were opportunities to be had with fracking that was going on in South Texas." "We're coming to the end of the bridge," Mr. Castro said. He said while he isn't calling for an immediate national ban on fracking, he supports local communities that do so and called for moving to cleaner sources. Mr. Castro won some of his strongest applause with the need to protect low-income communities from environmental damages.

ANSWER: no

REASON: The article does not provide a human personal story. It primarily focuses on Julián Castro's stance on fracking and environmental policies, without delving into any personal anecdotes or experiences.

SUMMARY:

EXAMPLE3:

ARTICLE: Greta Thunberg FAILS her 'zero emissions' promise after climate summit moves to Spain Greta Thunberg tried really hard to keep her "zero emissions" promise, but it has collapsed like all hypocritical left-wing plans because it was just too difficult. From the Hill:Greta Thunberg's plan for carbon-free travel was contradicted when her sailboat captain flew across the Atlantic Ocean to pick her up. Thunberg, 16, has been on a tour of North America to raise awareness for climate change.

ANSWER: no

REASON: The article does not provide a human story. the article criticizes Greta Thunberg for not being able to keep her promise of zero emissions, but it does not mention or provide any private anecdote or emotional language that would trigger feelings.

SUMMARY: [/EX]

[TASK] ARTICLE:

{article} [/TASK]

Figure 16: Mixtral prompt for labelling articles with HI-stories. {article} represents a placeholder for each of the article samples. Everything else is part of the prompt.

Prompt for filtering climate change articles

[*INST*] Climate change refers to the long-term warming of the planet, which is primarily caused by the increasing levels of greenhouse gases in the Earth's atmosphere resulting from human activities such as burning fossil fuels, whether deforestation, and land-use changes. You will be given a news article and your task is to predict the article is primarily about climate change or not. Your output should be ANSWER: yes, or ANSWER: no. Please provide your REASON justifying your answer. Take into account the criteria below to assess each article. Climate change relevant information includes the science behind climate change, its impacts, relevant political discussion, climate-related disasters, ways to address climate change, or related topics. Answer yes, if the article directly addresses one or more of these topics, as its other primary focus. Answer no if the article does not mention climate change, or only indirectly addresses it (e.g., one or two mentions of climate change in a poem or a person's story, who happens to be also climate change activist). Please mention in your REASON how much of the article is about climate change. Here are four examples, two positive and two negatives. [*INST*]

[*EX*] EXAMPLE1:

ARTICLE: Trump signs executive orders making it tougher for states to block pipeline construction President Trump on Wednesday signed two executive orders that will make it harder for states to block the construction of oil and gas pipelines and other energy projects due to environmental concerns. Coming on the heels of officials in Washington state and New York using permitting processes to stop new energy projects in recent years – and at the urging of some industry leaders – Trump's executive order will speed up the construction of oil and gas pipelines across the country." My action today will cut through ... [cut-off due to space]

ANSWER: yes

REASON: The article discusses climate-related policies, specifically the executive orders aimed at streamlining the permitting process for oil and gas pipelines, which is a climate change-related policy issue. The article explores the implications of these policies on the environment and the energy sector, which is closely tied to climate change. About 70 – 80% of the article is about climate change.

EXAMPLE2:

ARTICLE: Activists 'Shut Down D.C.' To Demand More Action On Climate Crisis Mark Wilson Getty Images Climate change protesters block traffic during a protest to shut down Washington, D.C., on Sept. 23, 2019. Activists demanding a greater global response to climate change disrupted Monday morning commutes with protests in Washington, D.C. Ahead of the day's United Nations Climate Action Summit in New York, activists participated in "Shut Down D.C.," a coordinated act of civil disobedience designed to draw attention to the global climate crisis. ... [cut-off due to space]

ANSWER: yes

REASON: The article is clearly focused on the climate crisis and the actions people are taking to demand greater action from governments and policymakers to address it. About 90 – 100% of the article is about climate change.

EXAMPLE3:

ARTICLE: India, March 19 – ChatGPT has been making headlines since its launch. Be it the chatbot acing simulated exams or giving "spot on" answer to being asked about Guinness World Records, social media is filled with posts about interactions with the AI. Just like this post where a man shared how ChatGPT wrote an "extraordinary" poem on climate change in the style of William Shakespeare. Twitter user Dan Miller, shared a Twitter thread showing how the AI chatbot wrote a poem for him. "I asked #ChatGPT4 to explain #ClimateChange in the voice of #Shakespeare. The result is quite extraordinary," he tweeted. In the next few tweets, he added a few lines from the poem generated by ChatGPT. ... [cut-off due to space]

ANSWER: no

REASON: The article is related to climate change, but only tangentially. While climate change is mentioned as the subject of the poem, the article does not provide any information or insights about climate change itself. Therefore, I would not classify this article as being primarily about climate change. It is more of a technology/human-interest article that happens to mention climate change as a topic. Only about 36% of the article is about climate change, and it is only indirectly addressed through the poem. [*EX*]

[*TASK*] ARTICLE:

{article} [*TASK*]

Figure 17: Prompt to Mixtral for selecting articles with climate change as the primary topic.