FRAPPE: FRAming, Persuasion, and Propaganda Explorer

Ahmed Sajwani[1,5*], **Alaa El Setohy**[10,5*], **Ali Mekky**[2,5*], **Diana Turmakhan**[3,5*], **Lara Hassan**[2,5*], **Mohamed El Zeftawy**[2,5*], **Omar El Herraoui**[4,5*], **Osama Mohammed Afzal**[5], **Qisheng Liao**[5], **Tarek Mahmoud**[5,8], **Zain Muhammad Mujahid**[5], **Muhammad Umar Salman**[5], **Muhammad Arslan Manzoor**[5], **Massa Baali** [5], **Jakub Piskorski**[6], **Nicolas Stefanovitch**[9], **Giovanni Da San Martino**[7], **Preslav Nakov**[5]

[1]Khalifa University of Science and Technology, [2]Alexandria University, [3]Nazarbayev University, [4]NYU Abu Dhabi, [5]MBZUAI, [6]Polish Academy of Sciences, [7]University of Padova, [8]Presight
[9]European Commission Joint Research Centre, [10]Egypt Japan University of Science and Technology

Abstract

The abundance of news sources and the urgent demand for reliable information have led to serious concerns about the threat of misleading information. In this paper, we present FRAPPE, a Framing, Persuasion, and Propaganda Explorer system. FRAPPE goes beyond conventional news analysis of articles and unveils the intricate linguistic techniques used to shape readers' opinions and emotions. Our system allows users not only to analyze individual articles for their genre, framings, and use of persuasion techniques, but also to draw comparisons between the strategies of persuasion and framing adopted by a diverse pool of news outlets and countries across multiple languages for different topics, thus providing a comprehensive understanding of how information is presented and manipulated. FRAPPE[1] is publicly accessible at https://frappe.streamlit.app/ and a video explaining our system is available at https://www.youtube.com/watch?v=3RlTfSVnZmk

1 Introduction

As the digital age ushers in an era of unparalleled connectivity and information dissemination, the transition towards online news media has brought several changes in the way the news is produced, consumed, and distributed. Although online news offers advantages such as increased accessibility and reduced cost of publishing, it has also brought several challenges such as the potential reinforcement of biases and propaganda. Consequently, analyzing news articles and offering an in-depth understanding beyond surface-level text has become pivotal for addressing these challenges. To get a global picture, there is also a need to analyze and to compare entire news media as well as the news landscape in different countries around a given topic.

A number of manual fact-checking initiatives have been launched. For example, Media Bias/Fact Check[2] and NewsGuard[3] offer assessments of the biases and the factuality of reporting of entire news outlets. There have also been a number of initiatives for fact-checking individual claims such as FactCheck,[4] PolitiFact,[5] and Snopes,[6] among many others. However, they all require tedious manual work by human fact-checkers, which does not scale and cannot cope with the volume of disinformation online. Thus, automation has been proposed as a possible alternative. Recently, it has been further realized that it is important to focus not only on factuality, but also on the way the message is conveyed, e.g., using subjectivity/humor, framing, and persuasion techniques.

As a parallel research line, a variety of research systems have been developed for the purpose of news analysis across several dimensions. For example, the Prta system (Da San Martino et al., 2020) allows users to explore the use of 18 propaganda techniques in news articles. Another tool, NewsLens (Laban and Hearst, 2017), focuses on constructing cohesive narrative threads that span several years and articles from approximately 20 news sources. Yet another system, Tanbih (Zhang et al., 2019), offers a profiling mechanism that covers a substantial yet confined collection of a few thousand media sources, which it profiles for factuality and bias of reporting. We can also mention NewsScan (Kevin et al., 2018), which is a plugin to profile news articles on the basis of specific labels providing information about the lexical properties (ease of reading), as well as some intrinsic properties (sentiment score, political bias), referred to as *nutrition labels*.

*Equal contribution.

[1]This work was done during a summer internship at the NLP department, MBZUAI.

[2]https://mediabiasfactcheck.com/
[3]https://www.newsguardtech.com/
[4]https://www.factcheck.org/
[5]https://www.politifact.com/
[6]https://www.snopes.com/

Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics System Demonstrations, pages 207–213

However, there is a lack of publicly accessible tools with the capability not only to analyze articles, but also to systematically explore and to draw comparisons between the strategies of persuasion and framing adopted by a diverse pool of news outlets and countries across many languages for a specific topic. With the aim to bridge this gap, we developed FRAPPE (FRAming, Persuasion, and Propaganda Explorer), an online news analysis platform that operates in a multilingual setup and provides two main functionalities: on-the-fly analysis for individual articles, and a user-friendly, interactive, and insightful Web interface for analyzing a database of 2M+ articles from 8k+ different media sources around the globe, in a variety of languages, covering two main topics: (*i*) the Russia–Ukraine conflict and (*ii*) climate change. This enables users to gather insights about framing, persuasion, and propaganda at an aggregate level, by news outlet and by country, and also to compare different news outlets and different countries.

2 Data and Models

2.1 Data

Our models were trained on a multilingual multi-faceted dataset of news articles from SemEval-2023 task 3 on "Detecting the Genre, the Framing, and the Persuasion Techniques in Online News in a Multi-Lingual Setup" (Piskorski et al., 2023b). The dataset consists of 1,612 articles covering news on current topics of public interest in six European languages (English, French, German, Italian, Polish, and Russian), with more than 37k annotated spans. Each news article was annotated for genre, framings, and persuasion techniques.

2.2 Model

Our system has trained on models for the three subtasks of SemEval-2023 task 3, which we briefly describe below.

2.2.1 Genre

For genre classification, our objective is to define the intended nature of an article, distinguishing between opinion pieces, objective news reporting, and satire. This categorization follows a multi-class annotation scheme applied at the article level.

We used an advanced transformer-based pre-trained multilingual language model, XLM-RoBERTa (Conneau et al., 2020), and in particular its base-sized model, which has 279M parameters.

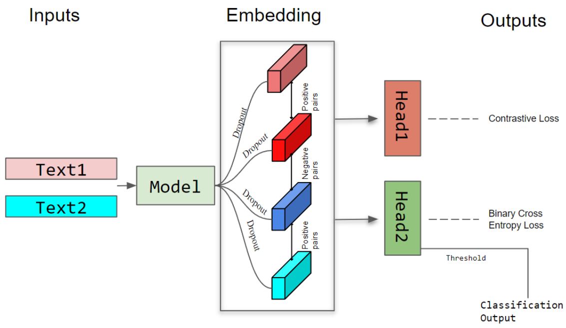

Figure 1: **The architecture of our system for framing and propaganda.** Each input generates two views of representations using dropout. The representations generated by the same text are positive pairs. The contrastive loss and the binary cross entropy loss are calculated separately by two heads and the classification result is calculated after applying a threshold.

This model has excellent transfer learning capabilities. To tailor it specifically to our genre classification task, we fine-tuned the model's last layer by adding a fully connected linear layer with three units (representing the three classes) and applying a softmax activation function.

The training dataset exhibits class imbalance, wherein the presence of satire articles is significantly limited in comparison to the abundance of opinion ones. Consequently, this introduced bias in the model's performance. To mitigate this issue, we used a focal loss function (Mukhoti et al., 2020), a specialized variant of the standard cross-entropy loss, which assigns higher weights to misclassified minority class examples, thereby emphasizing their significance during training and improving the model's ability to handle class imbalance.

Table 1 shows the performance of our model on the test sets for the six languages, with comparison against the baselines and the best systems at SemEval-2023 task 3 subtask A.

Language	Baseline	Our Sys	Best Sys
English	0.288	0.533	0.784
French	0.568	0.686	0.835
German	0.630	0.726	0.819
Italian	0.389	0.621	0.768
Polish	0.490	0.682	0.785
Russian	0.398	0.641	0.755

Table 1: **Genre analysis:** Performance (macro-F1 score) of our system compared to the baselines and to the best systems for the six languages on the test data of SemEval-2023 task 3 subtask A.

Language	Baseline	Our Sys	Best Sys
English	0.349	0.562	0.578
French	0.328	**0.552**	0.552
German	0.487	**0.711**	0.711
Italian	0.485	**0.617**	0.617
Polish	0.593	**0.673**	0.673
Russian	0.229	**0.449**	0.449

Table 2: **Framing analysis:** Performance (macro-F1 score) of our system compared to the baselines and to the best systems for the six languages on the test data of SemEval-2023 task 3 subtask B.

Language	Baseline	Our Sys	Best Sys
English	0.195	0.329	0.375
French	0.240	0.436	0.468
German	0.316	**0.529**	0.529
Italian	0.397	0.548	0.550
Polish	0.179	0.406	0.430
Russian	0.207	**0.395**	0.395

Table 3: **Propaganda analysis:** Performance (macro-F1 score) for our system compared to the baselines and the best systems for the six languages on the test data of SemEval-2023 task 3 subtask C.

2.2.2 Framing

For news framing, we trained our models on the data from SemEval-2023 task 3 subtask B (Piskorski et al., 2023a). This is a challenging task and it is more nuanced than mere topic classification (Card et al., 2015), e.g., while the topic of a news article may be COVID-19, the framing could be from an economic, political, or/and health perspective(s). The task can be formulated as a multilabel text classification problem with fourteen possible labels: *Economic; Capacity and Resources; Morality; Fairness and Equality; Legality, Constitutionality, and Jurisprudence; Policy Prescription and Evaluation; Crime and Punishment; Security and Defense; Health and Safety; Quality of Life; Cultural Identity; Public Opinion; Political; External Regulation and Reputation*. The performance of our model on the test data is shown in Table 2.

Our model's architecture (Liao et al., 2023) is illustrated in Figure 1, featuring two heads: one for contrastive loss and another one for binary entropy loss. During training, the focus is on optimizing the contrastive loss, which is achieved by creating two positive examples from each training example using dropout. Despite the dropout module altering the embeddings of the same training sample, they are still considered positive examples, and their distances are brought closer together. In the context of being a multi-label classification model, any other examples within the same batch are deemed positive only if they have exactly the same target classes. Otherwise, they are treated as negative examples, and the distance between them is pushed apart. Moreover, the loss of the negative examples is weighted, with higher weights assigned to examples containing more diverse classes. Finally, we transform the multi-label classification task into 14 binary classification ones: one for each frame label.

As a result, for each input, there are multiple logits corresponding to each target class. The classification decision is made by predicting that a given example belongs to a target class if the value of the corresponding logit exceeds a specific threshold.

2.2.3 Propaganda

In our propaganda model training, we used data from SemEval-2023 task 3 subtask C and the same model as before. The setup is comparable to the second subtask, with the distinction that the predictions are made for each sentence of the article, rather than for the entire article as a whole. Thus, the model generates predictions at the sentence level. After obtaining predictions for each individual sentence of the article, we combined these predictions to form the final multi-label prediction. This allows us to make comprehensive predictions that take into account the information from each sentence within the article, leading to a more nuanced and context-aware classification of propaganda elements in the text. There are 23 persuasion techniques in total: *appeal to authority; appeal to popularity; appeal to values; appeal to fear/prejudices; flag waving; causal oversimplification; false dilemma or no choice; consequential oversimplification; straw man; red herring; whataboutism; slogans; appeal to time; conversation killer; loaded language; repetition; exaggeration or minimisation' obfuscation - vagueness or confusion; name calling or labeling; doubt; guilt by association; appeal to hypocrisy; and questioning the reputation*. These 23 techniques are grouped into 6 coarse-grained categories: *Attack on Reputation, Justification, Simplification, Distraction, Calls, Manipulative Wording*. The performance of our model on the test data is shown in Table 3.

3 System Architecture

FRAPPE is a comprehensive system with two subsystems. The first one enables users to analyze differences in framings and propaganda techniques across countries and media sources. The second one allows users to explore the genre, the framings, and the propaganda techniques used in an individual custom article, which can be analyzed on the fly. We implemented both subsystems using Streamlit, a free and open-source framework for building and sharing machine learning and data science Web applications.

News Media Explorer We applied our custom-developed models, focusing on framing and persuasion techniques, to a dataset of 2,281,254 articles about the Russia-Ukraine conflict, sourced from 8,318 media outlets across 196 countries. These articles were processed and indexed on our platform, utilizing the models converted to ONNX format and deployed on NVIDIA Triton for enhanced inference speed, powered by two NVIDIA RTX 4090 GPUs. The aggregated results can be explored using the demo, enabling users to conduct customized analysis based on their preferences.

Custom Article Analyzer Inference is done using the models uploaded on the server, and the user is prompted to enter either a news article's URL or the text of an article. In the former case, we use Trafilatura, a Python package and command-line tool, to gather and to extract the article text from the URL (Barbaresi, 2021). After the predictions are done, the genre of the article is displayed, as well as the distribution of the framings and the propaganda techniques it uses.

4 Interface

Our system interface is designed with user friendliness and accessibility in mind, and it offers an immersive journey into the analysis of the two collections of articles we currently have loaded. The interface is divided into two engaging subsections. The News Media Explorer allows the user to explore the analysis of a collection of articles conveniently aggregated both by their source and country of origin, while the Custom Article Analysis offers real-time processing capability for an individual article.

4.1 The News Media Explorer

Setting sail on this analytical journey, users are welcomed by five interactive pages. Each page offers a unique perspective visualizing the following:

1. Framings and persuasion techniques for countries;

2. Framings for countries and sources;

3. Persuasion techniques for countries and sources;

4. Coarse-grained persuasion techniques for countries and sources;

5. Rhetorical propaganda techniques (ethos, pathos, logos) for countries and sources.

The user can choose a country from a list of 186 countries, which are shown sorted by total number of articles. From here, she is guided through a series of visualizations, each revealing detailed insights into the distribution of framings and persuasion techniques across countries and sources.

Figures 2 and 3 are the first landmarks on this journey, visualizing the distribution of framings by country and by source, respectively. The y-axis lists the countries/sources in descending order of articles, while the x-axis gives the percentage of each framing in each country/source. The legend, a colorful tapestry of framings, aids in understanding the chart.

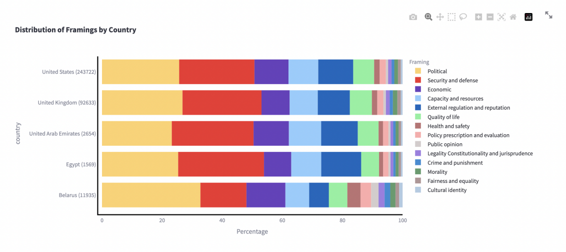

Figure 2: Visualization of framings by country.

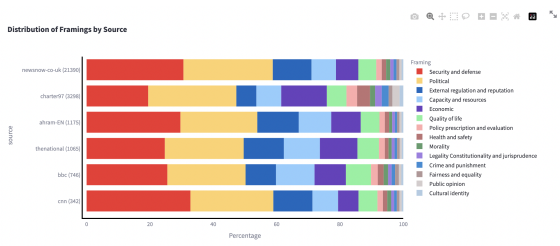

Figure 3: Visualization of framings by source.

The percentage of framings in each country/source (F) is calculated using the following formula:

$$F = \frac{f_{s,c}}{f_{T,c}} \times 100 \qquad (1)$$

where $f_{s,c}$ is the frequency of a specific framing within all the articles in a given country/source, and $f_{T,c}$ is the total number of articles for that country/source.

Next, Figures 4 and 5 unfold the story of persuasion techniques by country and source, respectively. Note that we detect the persuasion techniques at the sentence level, considering multiple instances within a single article. Thus, the percentage of each persuasion technique (P) in each country is calculated as follows:

$$P = \frac{F_{t,c}}{F_{T,c}} \times 100 \qquad (2)$$

where $F_{t,c}$ is the frequency of a specific persuasion technique in all articles for a given country/news source, and $F_{T,c}$ is the total frequency of all persuasion techniques in all articles from that country/source in our collection.

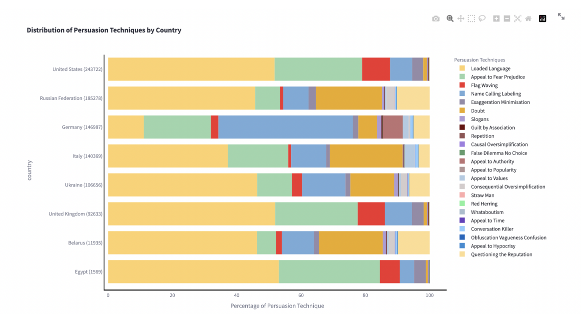

Figure 4: Visualization of persuasion techniques by country.

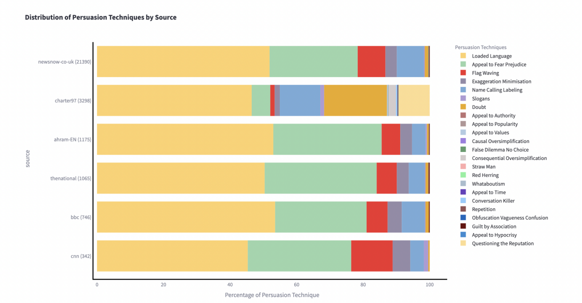

Figure 5: Visualization of persuasion techniques by news medium.

As the journey progresses, a graph displaying the number of articles over time for each country/source emerges, as shown in Figure 6, shedding light on the evolution of media trends.

The final stop within this subsection is an interactive pie chart that presents a global view of all framings in all the selected countries and their respective continents, as shown in Figure 7. This interactive feature invites users to click on parts of the pie chart to collapse or to expand sections, offering an exploratory experience.

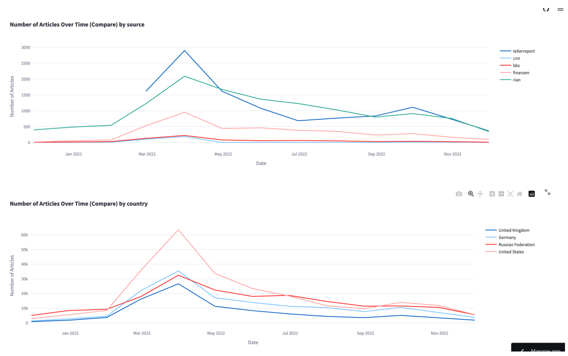

Figure 6: Visualization of the number of articles over time by country and by media source.

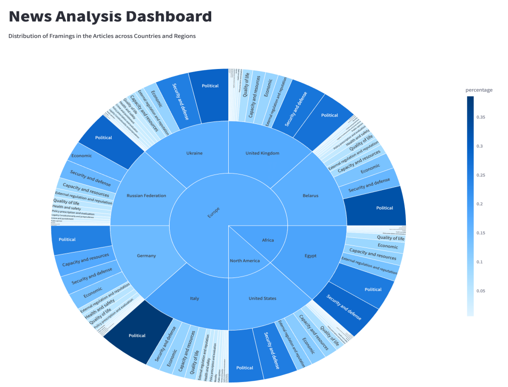

Figure 7: Pie chart of framings by country and continent.

The user can choose to explore the persuasion techniques at two coursers level (as opposed to the original 23 fine-grained techniques) by media and by country. Figure 8 shows the distribution of coarse-grained persuasion techniques across countries, allowing for a higher-level understanding of the prominent persuasion styles used. Figure 9 presents the use of rhetorical techniques such as ethos, pathos, and logos by country, highlighting the variations in strategy and deepening our comprehension of the media persuasion dynamics.

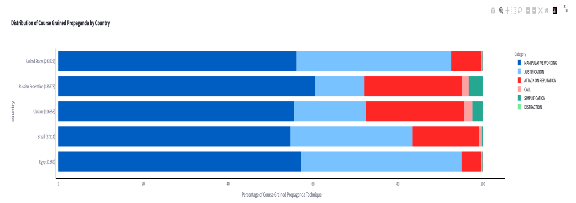

Figure 8: Visualization of coarse-grained persuasion techniques by country.

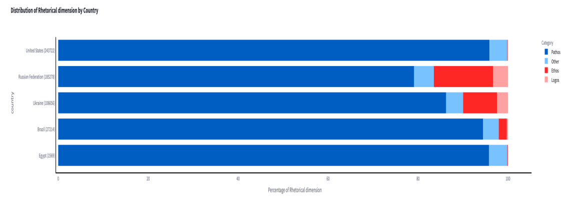

Figure 9: Visualization of rhetorical persuasion techniques by country.

4.2 Custom Article Analysis

The second part of the system, Custom Article Analysis, allows users to delve into a single article of their choice. Users can input their custom article either by entering its URL or by pasting its text into a text box. Upon submission, the system presents three enlightening visualizations, as shown in Figures 10, 11, and 12:

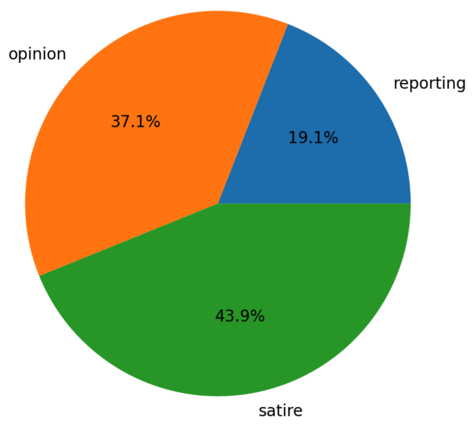

Figure 10: Classification of a custom article as reporting, opinion, or satire.

1. A pie chart unveiling the article's classification as satire, reporting, or opinion. This offers clarity about the article's genre, as shown in Figure 10.

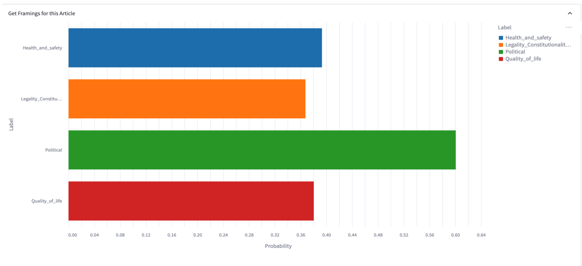

Figure 11: Framings graph for a custom article.

2. A bar chart revealing the framings in the article, offering a deeper understanding of the article's perspectives, as shown in Figure 11.

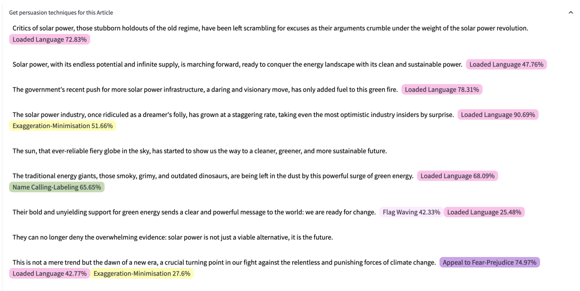

Figure 12: Persuasion techniques for a custom article.

3. A visualization of the persuasion techniques in the article, offering granular insight into the article's rhetoric. This tool is a valuable asset for anyone studying or practicing journalism, as shown in Figure 12.

5 Conclusion and Future Work

We presented FRAPPE, a system for analysis of the news in terms of framing, persuasion, and propaganda techniques, which can be used to tackle the challenge of misleading information in news articles and news outlets. It provides a deep understanding of news articles, allowing users to gain insights into genre, framing, and propaganda techniques, and to perform comparison of media and countries.

In future work, we aim to gather more data, to extend the analysis to other globally discussed topics, and to update our visualizations with more in-depth analysis that goes down to the individual article level. We further plan to create a database that will automatically save every article that was analyzed on the fly (in a separate collection). Finally, we aim to improve our models' accuracy, and to add additional models and analysis.

6 Limitations

We acknowledge certain limitations that we plan to address in future work. First, we need to expand the training data to cover more languages as part of training, in order to improve our models' accuracy when analyzing articles written in these languages (even though, thanks to the multilingual XLM-RoBERTa, we already cover 100 languages). Second, our system does not cover all aspects of analysis that an article can undergo; it currently only unveils genre, framings, and persuasion techniques. Finally, our database is currently limited to our 2M+ articles, and does not automatically reflect future events. We plan continuous data addition covering more topics and languages.

7 Ethics and Broader Impact

One of the foremost ethical considerations is ensuring the transparency and the unbiased analysis in FRAPPE. Users should be aware that our models use neural networks, and as such, they lack explainability. Another warning is that, despite our intent, due to article selection biases, FRAPPE might be favoring some political or social standpoints.

Finally, FRAPPE has the potential to influence the way news articles are perceived and consumed, and journalists may become more aware of the language they use and its potential impact on readers.

References

Adrien Barbaresi. 2021. Trafilatura: A web scraping library and command-line tool for text discovery and extraction. In *Proceedings of the Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations*, ACL-IJCNLP '21, pages 122–131.

Dallas Card, Amber E. Boydstun, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2015. The media frames corpus: Annotations of frames across issues. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, NAACL-HLT'15, pages 438–444, Beijing, China.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, ACL '20, pages 8440–8451.

Giovanni Da San Martino, Shaden Shaar, Yifan Zhang, Seunghak Yu, Alberto Barrón-Cedeño, and Preslav Nakov. 2020. Prta: A system to support the analysis of propaganda techniques in the news. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, ACL '20, pages 287–293.

Vincentius Kevin, Birte Högden, Claudia Schwenger, Ali Şahan, Neelu Madan, Piush Aggarwal, Anusha Bangaru, Farid Muradov, and Ahmet Aker. 2018. Information nutrition labels: A plugin for online news evaluation. In *Proceedings of the First Workshop on Fact Extraction and VERification*, FEVER '18, pages 28–33, Brussels, Belgium.

Philippe Laban and Marti Hearst. 2017. newsLens: building and visualizing long-ranging news stories. In *Proceedings of the Events and Stories in the News Workshop*, pages 1–9, Vancouver, Canada.

Qisheng Liao, Meiting Lai, and Preslav Nakov. 2023. MarsEclipse at SemEval-2023 task 3: Multi-lingual and multi-label framing detection with contrastive learning. In *Proceedings of the the 17th International Workshop on Semantic Evaluation*, SemEval '23, pages 83–87, Toronto, Canada.

Jishnu Mukhoti, Viveka Kulharia, Amartya Sanyal, Stuart Golodetz, Philip Torr, and Puneet Dokania. 2020. Calibrating deep neural networks using focal loss. *Advances in Neural Information Processing Systems*, 33:15288–15299.

Jakub Piskorski, Nicolas Stefanovitch, Giovanni Da San Martino, and Preslav Nakov. 2023a. SemEval-2023 task 3: Detecting the category, the framing, and the persuasion techniques in online news in a multi-lingual setup. In *Proceedings of the 17th International Workshop on Semantic Evaluation*, SemEval'23, pages 2343–2361, Toronto, Canada.

Jakub Piskorski, Nicolas Stefanovitch, Nikolaos Nikolaidis, Giovanni Da San Martino, and Preslav Nakov. 2023b. Multilingual multifaceted understanding of online news in terms of genre, framing, and persuasion techniques. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, ACL '23, pages 3001–3022. Association for Computational Linguistics.

Yifan Zhang, Giovanni Da San Martino, Alberto Barrón-Cedeño, Salvatore Romeo, Jisun An, Haewoon Kwak, Todor Staykovski, Israa Jaradat, Georgi Karadzhov, Ramy Baly, Kareem Darwish, James Glass, and Preslav Nakov. 2019. Tanbih: Get to know what you are reading. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing: System Demonstrations*, EMNLP-IJCNLP '19, pages 223–228, Hong Kong, China.