ESG Tendencies from News - Investigated by AI Trained by Human Intelligence

Li, Chao and Keeley, Alexander Ryota and Takeda, Shutaro and Seki, Daikichi and Managi, Shunsuke

Kyushu University

November 2024

ESG Tendencies from News

- Investigated by AI Trained by Human Intelligence

Abstract

We create a large language model with high accuracy to investigate the relatedness between 12 environmental, social, and governance (ESG) topics and more than 2 million news reports. The text match pre-trained transformer (TMPT) with 138,843,049 parameters is built to probe whether and how much a news record is connected to a specific topic of interest. The TMPT, based on the transformer structure and a pre-trained model, is an artificial intelligence model trained by more than 200,000 academic papers. The cross-validation result reveals that the TMPT's accuracy is 85.73%, which is excellent in zero-shot learning tasks. In addition, combined with sentiment analysis, our research monitors news attitudes and tones towards specific ESG topics daily from September 2021 to September 2023. The results indicate that the media is increasing discussion on social topics, while the news regarding environmental issues is reduced. Moreover, towards almost all topics, the attitudes are gradually becoming positive. This research highlights the temporal shifts in public perception regarding 12 key ESG issues: ESG has been incrementally accepted by the public. These insights are invaluable for policymakers, corporate leaders, and communities as they navigate sustainable decision-making.

Keywords:

ESG; News; Natural Language Processing; Pre-trained Transformer; Data Mining; Machine Learning

Introduction

Due to climate change and the economic recession after the COVID-19 pandemic, various environmental and social problems have gradually become prominent (Aburto, Tilstra, Floridi, & Dowd, 2022; Le Billon, Lujala, Singh, Culbert, & Kristoffersen, 2021; X. Li, Huang, Li, & Xu, 2022; Vo et al., 2023). Corporations, as the major social and economic participants, are desired to be more responsible in society (Cezarino, Liboni, Hunter, Pacheco, & Martins, 2022; Le, Vo, & Venkatesh, 2022; Pan, Abbas, Álvarez-Otero, Khan, & Cai, 2022; Weston & Nnadi, 2023). The United Nation's Principles for Responsible Investment report introduces an innovative and holistic framework for Environmental, Social, and Governance (ESG) to lead companies' value transformation from pursuing profits to balancing performance across environmental, social, and economic dimensions (Gibson Brandon, Glossner, Krueger, Matos, & Steffen, 2022; Litvinenko, Bowbrick, Naumov, & Zaitseva, 2022). To implement ESG, corporations shouldering corporate social responsibility (CSR) act as practitioners, governments and non-government organizations (NGOs) play as referees, and the public participates as supervisors. Generally, the corporations' attitudes towards ESG are reflected in their annual reports (Baier, Berninger, & Kiesel, 2020; Nobanee & Ellili, 2016). Additionally, governments and NGOs could regulate and guide corporations by making laws, regulations, and guidelines, which mirror referees' attitudes (He, Jing, & Chen, 2023). However, the public's attitude is hard to grasp since the public contains millions of individuals.

Media are invaluable resources for investigating change and development in the economy and society (Dang, Dang, Hoang, Nguyen, & Phan, 2020; Dang, Dang, Moshirian, Nguyen, & Zhang, 2019; Liu, Luo, & Lu, 2023). On the one hand, the contents disclosed by the media could offer clear patterns of status to the public to reduce information asymmetry (Ding, Appolloni, & Shahzad, 2022; Lerouge, Lema, & Arnaboldi, 2023). On the other hand, a barrage of news guides the direction of public opinion (Moore, Dahlke, & Hancock, 2023) and contributes to some important topics, including achieving sustainable development, mitigating climate change, facilitating CSR, among others (Anita, Shveta, Yadav Surendra, & Arvind, 2023; Yu, Liang, Liu, & Wang, 2023). Compared with governments and NGOs' documents and corporations' annual reports, the news is free of much background knowledge, concise, brief, and understandable. Different from the contents of social networking services, the news from major media is relatively reliable. Of course, the media also report governments' policies, NGOs' actions, and firms' performance. Previous studies declare the change in a certain industry or stock price by summarizing the sentiment of all available news in related scopes (Dang et al., 2020; Dang et al., 2019). Specifically, if the tone of an industry is filled with doom and gloom, then most investors will be cautious or avoid investing in it. Aggregating the tones of all news is a practical approach because market confidence is vital in the current economy, and news represents mainstream thinking of the market. In addition, the popularity of news on a certain topic can also promote and mirror people's perception of certain topics to a certain extent (Zhang, Pan, Yu, & Liu, 2022). Therefore, extensive and efficient analysis of media news is of great value in helping policy-makers formulate appropriate policies.

This study examines the temporal variations of the public's attitude and tone towards the ESG topics of interest based on the news. Our approach combines relevance analysis and sentiment analysis: specifically, relevance analysis filters out news related to a topic, and sentiment analysis determines the tone of these news articles. If the media has a long-term pessimistic attitude towards a topic, then there should be problems with this topic, and vice versa. Based on previous research and hot issues, we choose the

ESG topics. In environmental aspects, with soaring energy prices, increasing numbers of natural disasters, and a variety of health risks, greenhouse gas emissions and air pollution have been of broad concern for a long time (Geng et al., 2023; Lelieveld, Evans, Fnais, Giannadaki, & Pozzer, 2015; Wekhof & Houde, 2023). These hot spots are highly relevant to people's lives, e.g., globally $PM_{2.5}$ causes approximately 3.3 million premature deaths each year (Lelieveld et al., 2015). Moreover, in the past several years, international society has been paying more attention to human rights, especially whether there are human rights issues in the entire supply chain of multinational company products (Ullah, Adams, Adams, & Attah-Boakye, 2021). Based on this background, many issues concerning employment and labor conditions are hotly discussed in social aspects (Cohen, 2023; Fiaschi, Giuliani, Nieri, & Salvati, 2020; Gervais, Kleijn, Nold, & van der Voet, 2023; York et al., 2022). Affected by high inflation and U.S. monetary policies, ESG topics related to employment have gained more voice (Bennani, 2023; Gregory, 2022; Luu & Palczewski, 2018; Ng, Lye, Chan, Lim, & Lim, 2020). Basic human rights related to infrastructure have also received widespread attention from all sectors of society (Douthit, Kiv, Dwolatzky, & Biswas, 2015; Hurlbert, 2020; WHO, 2020). According to WHO's reports in 2019, 2.1 billion people lack safely managed drinking water mainly living in poor and undeveloped regions (WHO, 2020). As a basic survival need, drinking water is regarded as a human right attracting social attention (Hurlbert, 2020). Similarly, health care is also a basic human right issue with wide impacts on the society, but a significant regional gap exists even in developed countries, such as the U.S. (Douthit et al., 2015). Among ESG topics, the environmental and social aspects are mainly concerned by the general public, while the governance aspect is related to the company's leadership, executive pay, audits, internal controls, and shareholder rights majorly supervised by the investors and stakeholders (Cezarino et al., 2022; Keeley et al., 2022). This study aims to explore the changes in public attitudes rather than investors or stakeholders towards ESG-related topics through a large amount of news. Therefore, the topics of interest are principally from the environmental and social aspects.

The nexus between ESG topics and media coverage has increasingly attracted academic and practical attention, especially in the context of shifting public sentiments and regulatory landscapes post the COVID-19 pandemic. However, despite considerable research, there is a significant gap in quantitatively and systematically assessing how media coverage reflects and influences ESG-related public perceptions over time. Our study builds a novel artificial intelligence model with high accuracy to label available 2,164,153 news including more than 2 billion words published from September 1st, 2021, 00:00:00 (UTC) to September 1st, 2023, 00:00:00 (UTC) in the United States from 354 news sources. Furthermore, we use a publicly available sentiment model to analyze their tone. The results demonstrate a comprehensive overview of media discourse variations on ESG topics. By exploring the temporal shifts in media tones and attitudes towards ESG topics based on the relatedness analyses and sentiment analyses, this paper contributes novel insights into the dynamic interplay between media coverage and public ESG awareness. On the technical level, we also provide a paradigm for LLM-based news analysis. These contributions are vital for developing strategies that align corporate and public actions with sustainable development goals.

Materials and Methodology

Daily News

We grasp 2,830,314 news based on NewsData API published from September 1st, 2021, 00:00:00 (UTC) to September 1st, 2023, 00:00:00 (UTC) in the United States from 354 news sources, such as CNN, USAToday, NY Times, Washington Times, Washington Post, Forbes, among many others. The news could directly reflect the public's attitude towards a certain event or a specific topic and then also lead and guide society. However, some news is written in non-English in the dataset. To analyze the news by our LLM efficiently, we focus only on English news.

The usable texts include two types according to free availability because some news records are open-access, whereas others are not. For the open-access news, the data vendor, NewsData, could grasp the whole news content, while for the subscribed news, it might only obtain the news briefs depending on the data vendor's subscription list. The news briefs are generally the first several sentences of the news. Although their lengths are typically no longer than 100 words, we still investigate them because the most important information is normally arranged at the very beginning of the news records. Hence, in this way, we could judge whether the subscripted news records are related to a particular topic and analyze their sentiment. After washing the news data, 2,164,153 news records are available and examined. In the 730 days, we could obtain 2964.593 valid news records per day. **Figure 1** illustrates the bar chart of daily news count.

ESG-related Topics

With the introduction of ESG, the public and society gradually focus more closely on the environmental and social impacts of the corporations' actions (Aburto et al., 2022; Le Billon et al., 2021; X. Li et al., 2022; Vo et al., 2023). According to the recent research on ESG, in total, 12 topics are considered and analyzed. In the environmental aspect, "energy usage", "greenhouse gas", and "air pollution" are widely investigated. Additionally, in the social aspect, "forced labor", "freedom of association, collective bargaining, right to strike", "unemployment", "poverty and inequality", "corruption", "indigenous rights", "excessive working time", "access to an improved source of drinking water", and "access to health care" are taken in account. It should be noted that we shorten some topics. Specifically, the literal length of the topic is longer than our limit for the keyword in our analysis. The model only reads the first six tokens of the keywords. If longer than this requirement, the tokens from the seventh position would be ignored. To maintain the critical information on this topic, we replace it with "association collective bargaining strike".

Natural Language Processing of News

The public's attitude towards a topic should be positive if most news related to a topic of interest conveys cheerful tones. In this way, every available news should be scored objectively. The news scoring equation is as follows:

$$NS_{ijt} = NR_{ijt} \times NPD_{it} \times NPS_{it} \tag{1}$$

where NS_{ijt} represents the news score (NS) towards a certain topic j of the ith news published on the t day, NR_{ijt} represents the relatedness between the ith news published on the t day and the j topic among 12 listed topics, NPD_{it} represents a polarity of the ith news published on the t day, and NPS_{it} represents the probability of the ith news published on the t day labeled as a certain polarity. NR_{ijt} is a value ranging from 0 to 1, where 0 represents that the news is unrelated to the topic j whereas 1 represents that they are highly related. The news tone polarities are three types, precisely, positive, neutral, and negative. In the equation, positive, neutral, and negative polarities are valued as 1, 0, and -1, respectively. NPS_{ijt} is also the probability of the NPD_{it} from 0 to 1. The calculations of NR_{ijt}, NPD_{it}, and NPS_{it} are depicted later.

Text Match Pre-trained Transformer (TMPT)

We adopt a similar strategy to probe the variations of ESG topics based on available news as the previous studies (Dang et al., 2020; Dang et al., 2019; De Giuli, Grechi, & Tanda, 2023; Lerouge et al., 2023). However, we need a novel labeling method to judge whether the news is related to the selected ESG topics. Previous studies mainly focus on a specific industry or a certain entity. They mostly use the named entity recognition (NER) method to label the news (Dang et al., 2020; Dang et al., 2019). Specifically, the NER method needs the news to explicitly include the entity names (Marrero, Urbano, Sánchez-Cuadrado, Morato, & Gómez-Berbís, 2013). If a report is definitely linked to an entity, the entity's information, such as industry group, country, and industry sector, is conveyed to the labeled news. For our task, the NER method is not suitable. For instance, the news about "greenhouse gas" might only mention its synonyms rather than notice it directly. If we label the news by judging whether "greenhouse gas" appears in the news, the performance would be poor. The growing recognition and widespread application of generative pre-trained transformers (GPT) (Brown et al., 2020; Radford et al., 2019), like ChatGPT, suggest that large-scale models in the field of natural language processing (NLP) are steadily making inroads into diverse industries. To solve this issue, we create a new large language model (LLM), text match pre-trained transformers (TMPT), to probe the relatedness between news and the selected ESG topics. The TMPT is a cutting-edge artificial intelligence stemming from human intelligence. It aims to judge whether the news and topic are matched for "TM". To reduce the training workload, we transplant the embedding layer from a pre-trained model, Bidirectional Encoder Representations from Transformers (BERT) (Devlin, Chang, Lee, & Toutanova, 2018), which includes 109,482,240 parameters. This is the reason for "P". This model is based on the state-of-the-art machine learning technology, transformer, for "T" (Vaswani et al., 2017). The transformer block layer could review all input contents and pay more attention to the "core" part in the correct location. Because all the texts are learned during training, the synonyms in the abstracts are recognized and remembered by the TMPT. Therefore, this LLM could solve the issue smoothly.

The TMPT model is artificial intelligence stemming from numerous human intelligences, since we use 200,000 academic papers to train TMPT's 138,843,049 parameters. It is designed to quantitatively investigate the relatedness between a piece of news and a certain topic. The model could be abstracted as follows:

$$NR_{ijt} = R(News_{it}, Topic_j) \tag{2}$$

where R represents an estimation function based on our TMPT model, $News_{it}$ represents the ith news published on the t day, and $Topic_j$ represents the topic j of interest. It should be emphasized that the TMPT is a general model. In other words, this model could be used for other topics by changing $Topic_j$.

Training Dataset

Technically, the TMPT is a Siamese network. Specifically, the TMPT needs two inputs: the first one is a paragraph, and the second one is a keyword. To train the TMPT, we must use paragraphs with exact, reliable, informative, and sensible labels. The academic articles are such resources. Academic papers are normally written by educated researchers, and concise and informative abstracts and keywords are mandatory. The keywords could be regarded as the labels of the abstracts. Hence, the academic articles' abstracts and keywords are ideal for training the TMPT model.

To build the training dataset, in total, we download more than 200,000 articles with available abstracts and keywords. They are from Elsevier journals with high impact factors related to business, environment, society, and other similar scopes, such as Journal of Cleaner Production, Science of the Total Environment, Sustainable Cities and Society, among many others. We use one abstract and one keyword to make one record. If a record's abstract and keyword are from the same article, the record would be labeled as true, and its output value is 1. Simply, for the true records, the keywords are essentially and highly related to the abstracts. If they are not from the same article, the record is considered false, with 0 as its output value. Since an article generally has several keywords and we input one abstract and one keyword each time, the number of true records in the training dataset is larger than the number of downloaded articles. We employ the BERT tokenizer to encode abstracts and keywords. The BERT tokenizer dictionary does not include any customized abbreviations (Devlin et al., 2018). If the keywords are not widely used abbreviations, they should be dropped during the training dataset creation stage. More than 1,000,000 true records are created based on 200,000 articles. The TMPT model works for a binary classification task per se. To train the model effectively, the training dataset should be balanced, -i.e., the numbers of true and false records should be the same. We take the keywords from another article to replace the keywords used in the true records. To guarantee that our false abstracts and keywords are inconsistent, the “fake” keywords should not be the same as true words, a low cosine similarity between the “fake” and true keywords based on the BERT (Devlin et al., 2018), and from a different journal. The data size of the training dataset is approximately 2 million records.

The TMPT model’s inputs are two sequences of tokens rather than the texts of abstracts and keywords. Hence, before training the model, we need to encode the abstracts and keywords by the BERT tokenizer. In the encoding step, two processing strategies are critical, namely sequence length and padding. The end-to-end NLP models normally have a fixed length of input sequences. The abstracts of academic articles generally have 200 words, and they will be converted into approximately 300 tokens because the BERT uses WordPiece tokenization (Devlin et al., 2018), and every symbol is also encoded as a single token. The WordPiece tokenization breaks down words into smaller pieces, e.g., “running” is split into two parts, “run” and “##ning”, and encoded, respectively (Devlin et al., 2018). If the sequence length is significantly longer than academic abstracts, the model might be too large and cost more time to train and use. On the other hand, a short sequence length leads to ignoring too much important information. In this way, we carefully set the length as 300 tokens. For a similar reason, we set the keyword input length as six tokens, considering that the academic keywords include 2 to 5 words. To maintain every sequence the same length, we must adapt padding and truncation strategy. In this study, we use post-padding, that is, adding padding tokens to the end of the sequence if the encoded text length is lower than the defined sequence length. Additionally, we truncate the end of the encoded texts if its length is longer than the pre-defined value. To sum up, all abstracts and keywords would be encoded into 300-integer and 6-integer sequences, respectively.

The Architecture of the TMPT Network

Figure 2 illustrates the architecture of the TMPT model. The functions and necessary details of each layer are noted in the figure. The input layers are to instantiate the tensors of the encoded abstract and keyword. To further process the input tensors, embedding layers are employed to reduce the dimension of the instance (Devlin et al., 2018; Mikolov, Sutskever, Chen, Corrado, & Dean, 2013). Naturally, each element of the input sequences is categorical rather than numerical. Even though they are integer numbers, they are not additive. They should be rewritten into matrixes based on the one-hot encoding. In our case, each element should be transmuted in a one-hot vector with 30,522 elements. According to this rule, the input paragraph and keyword tensors are 300×30522 and 6×30522 matrixes. The sparse input matrixes often worsen the LLM's performance. Embedding technology directly solves this issue by reprojecting the words to a relatively lower-dimensional space (Mikolov et al., 2013). Furthermore, the synonyms are always close in the lower-dimensional space to reduce information loss. The embedding layer has the most parameters in models generally. To reduce the training time and improve the model's performance, the engineer and researcher prefer to transplant the embedding layer from a pre-trained model to their LLMs. Therefore, we transplant the BERT embedding layer, which includes 109,482,240 parameters to reproject 30,522-element vectors into 768-dimension space, to our TMPT. Compared with the model size, our training data is less, so we do not further fine-tune the embedding layer parameters anymore. The transformer block layers contain a multi-head self-attention layer and dense layers to secure dimensions.

With the self-attention mechanism, the model is trained to give higher weights on the potential connection between parts of the input paragraphs and keywords (Vaswani et al., 2017). For the paragraph branch, the transformer block's self-attention is eight-head, while for the keyword branch, it is two-head. With more heads in the self-attention layer, the performance might increase, but the calculation time would also be longer. We choose current settings for the current model owing to the balance between the performance and time cost. The layers between transformer block layers and difference layer are to tidy the size of tensors to make them comparable. Both paragraphs and keywords are converted into 512-element tensors. We calculate the difference between two tensors and then use the dense layers to reduce the dimension to 1.

The relatedness between the input paragraph and the input keyword is the output of the TMPT model. The last dense layer's activation function is a sigmoid function, whose output is a probability value. The true-false threshold is 0.5. In other words, if the output is greater than 0.5, we regard the inputs as related, and vice versa. Moreover, the higher the model output value is, the more highly linked the inputs are. Since we just care about the related parts of the result, the relatedness function *R* based on our TMPT is expressed as follows:

$$NR_{ijt} = \begin{cases} \left[tmpt\left(News_{it}, Topic_j\right) - 0.5\right] \times 2, if tmpt\left(News_{it}, Topic_j\right) > 0.5 \\ 0, if tmpt\left(News_{it}, Topic_j\right) \leq 0.5 \end{cases} \tag{3}$$

where $tmpt$ represents our TMPT model, and $tmpt\left(News_{it}, Topic_j\right)$ represents the output of our model when the inputs are $News_{it}$ and $Topic_j$.

We use TensorFlow to build the TMPT network on a virtual machine with 8 NVIDIA A100 40GB GPU in the Google Cloud Platform. One hundred thousand records are taken as the validation dataset. The optimizer is Adam, and the learning rate is 0.0001. The batch is set to 2,048. We set the model to be trained 100 epochs, and the early stop strategy according to validation loss with 20 epochs patience is employed.

Statistical Indicators

We evaluate the model's performance based on the four statistical indicators, including accuracy, precision, recall, and F1-score. These four indicators summarize the status of the predicted results. The model could correctly and incorrectly predict the labels. For correct predictions, there are two situations, namely correctly predicted either positive or negative results. The correctly predicted positive and negative instances are called true positive (TP) and true negative (TN). In the same way, the incorrect positive and negative predictions are named false positive (FP) and false negative (FN). Accuracy is the ratio of correct predictions to all data:

$$Accuracy = \frac{TP + TF}{TP + TF + FP + FN} \tag{4}$$

Precision is the ratio of correct predictions to positive predictions:

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

Recall is the ratio of correct predictions to actual positive data:

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

F1-score is the harmonic mean of precision and recall:

$$F1score = \frac{2 \times Precision \times Recall}{(Precision + Recall)} \tag{7}$$

Precision could warn the researchers and engineer whether the model is apt to label everything as positive, while recall indicates whether the model tends to label all data as negative. The F1-score is a unified metric that combines precision with recall.

They are important when the dataset is imbalanced. In our case, these four metrics should be similar because our training dataset is balanced.

Sentiment Analysis

To reduce the cost, we download a pre-trained model, named distilroberta-finetuned-financial-news-sentiment-analysis (DFFNSA), from HuggingFace to analyze the news sentiment. This model is trained based on the financial_phrasebank dataset (Malo, Sinha, Korhonen, Wallenius, & Takala, 2014). The model's accuracy is 97.79% after five epochs of training. We choose this model for several reasons. First, initially, this model is trained based on financial news, which is close to our topic. Second, the max length of this model input is 512 tokens, significantly longer than Twitter sentiment or comments sentiment analysis models. This length is more suitable for our analysis. Third, this relatively light model includes only 82.1 million parameters. The lightweight model could dramatically decrease the computing time of sentiment analysis. The output of DFFNSA is a three-element vector. Each element is a probability for a label among "negative", "neutral", and "positive". A SoftMax function connects the three elements, -i.e., the sum of three probabilities is 1. The sentiment of the news is estimated as follows:

$$NSentiment_{it} = argmax(Pnegative_{it}, Pneutral_{it}, Ppositive_{it}) \tag{8}$$

where $NSentiment_{it}$ represents the sentiment of the ith news published on the t day, which should be one of "negative", "neutral", and "positive", the $argmax$ represents the function returning "argument of the maximum", $Pnegative_{it}$ represents the probability of labeling the ith news published on the t day as "negative", $Pneutral_{it}$

represents the probability of labeling the news as “neutral”, and $Ppositive_{it}$ represents the probability of labeling the news as “positive”. The NPD_{it} is calculated as follows:

$$NPD_{it} = \begin{cases} -1, if\ NSentiment_{it} == "negative" \\ 0, if\ NSentiment_{it} == "neutral" \\ 1, if\ NSentiment_{it} == "positive" \end{cases} \tag{9}$$

The NPS_{it} is estimated as follows:

$$NPS_{it} = max(Pnegative_{it}, Pneutral_{it}, Ppositive_{it}) \tag{10}$$

Daily Scores from News

To summarize the news analysis results and probe the variations of ESG tendencies extracted from the news effectively and efficiently, we construct several indicators. They are daily average relatedness based on all available news (DARA), daily average NS based on all available news (DANSA), daily average relatedness based on the related available news (DARR), daily related news count (DRNC), and daily average NS based on the related available news (DANSR). The DARA and DANSA cover all the available news, monitoring whether the topics of interest are buried under a barrage of other content. Moreover, the DARR and DANSR investigate how much the news concentrates on topics of interest.

It must be emphasized that DARA, DANSA, DARR, DRNC, and DANSR correspond to a certain topic j on a specific day t, referred to as $DARA_{jt}$, $DANSA_{jt}$, $DARR_{jt}$, $DRNC_{jt}$, and $DANSR_{jt}$, respectively. We regard the news with a non-zore relatedness score NR_{ijt} as the related news towards a certain topic j. The $DARA_{jt}$, $DANSA_{jt}$, $DARR_{jt}$, $DRNC_{jt}$, and $DANSR_{jt}$ are estimated as follows:

$$DARA_{jt} = average_{DARA}(NR_{ijt}) \tag{11}$$

where $average_{DARA}$ is to calculate the average news relatedness based on all available news by the day t and the topic j.

$$DANSA_{jt} = average_{DANSA}(NS_{ijt}) \tag{12}$$

where $average_{DANSA}$ is to calculate the average news score based on all available news by the day t and the topic j.

$$DARR_{jt} = average_{DARR}(NR_{ijt}|NR_{ijt} > 0) \tag{13}$$

where $average_{DARR}$ is to calculate the average news score based on the related news by the day t and the topic j.

$$DRNC_{jt} = count(NR_{ijt} > 0) \tag{14}$$

where $count$ is to count the news related to the topic j in the day t.

$$DANSR_{jt} = average_{DANSR}(NS_{ijt}|NR_{ijt} > 0) \tag{15}$$

where $average_{DANSR}$ is to calculate the average news score based on the related news by the day t and the topic j.

Temporal Tendency and Correlation Analysis

We analyze the temporal tendency of $DARA_{jt}$, $DANSA_{jt}$, $DARR_{jt}$, $DRNC_{jt}$, and $DANSR_{jt}$ based on Pearson's correlation coefficient. Specifically, we compute the correlation coefficient between the daily indicators and the order of dates. The values of the correlation coefficients range from -1 to 1. If the correlation equals to -1, the variables of interest are entirely negatively correlated, statistically. On the other hand, if the variables are completely positively associated, the correlation coefficient should be 1. If the variables are unrelated, the correlation coefficient would not be significant or be closer to 0. In our temporal tendency analysis, September 1st, 2021 is regarded as the beginning of our dataset, whose order of dates is 0.

Results

Performance of TMPT Model

The TMPT model is the core tool to analyze the news. Hence, the performance of the TMPT model is critical. We use 1,940,750 records to train the model and randomly take 100,000 records as the validation dataset. The validation dataset contains 50,856 positive records and 49,144 negative records. In total, our TMPT model includes 138,843,049 parameters. Among them, 29,360,809 parameters are trained by our dataset, while 109,482,240 parameters are transplanted from the BERT. After 33 epochs of training, the model reaches the best performance. It predicts 50,612 positive records and 49,388 negative records. In detail, TP is 43,599; FN is 7,257; FP is 7,013; and TN is 42,131. The accuracy of the model in cross-validation is 85.73%. Furthermore, the precision, recall, and f1-score are 86.14%, 85.73%, and 85.93%, respectively. In terms of zero-shot learning, TMPT's performance is excellent.

Daily Summary

Table 1 summarizes the statistics of DARAs. Among the 12 investigated topics, "corruption" has gained the most attention, while "indigenous rights" is the least discussed. **Figure 3** demonstrates the temporal variations of DARAs. At first glance, the indexes from the social aspects, especially "freedom of association, collective bargaining, right to strike", "poverty and inequality", and "indigenous rights", are gradually gaining more emphasis, whereas the relatedness of the indexes from the environmental aspects is decreasing. According to the correlations between the order of dates and the DARAs towards each topic, air pollution relatedness is significantly negatively associated with the order of dates. The air pollution relatedness's correlation coefficient is -0.619 (p value < 0.1%). The correlations between the order of dates and greenhouse gas and energy usage relatedness are significant, specifically -0.430 (p value < 0.1%) and -0.404 (p value < 0.1%), respectively. Statistically speaking, among all available news, gradually less news is related to environmental ESG topics in the US from September 1^{st}, 2021, to September 1^{st}, 2023. The positive correlation between the order of dates and the DARA of "poverty and inequality" is the highest, which is 0.703 (p value < 0.1%). Other DARAs of social indexes, including "forced labor", "freedom of association, collective bargaining, right to strike", "unemployment", "corruption", "indigenous rights", and "access to health care", are also positively correlated with the order of dates, where the correlation coefficients are 0.353 (p value < 0.1%), 0.679 (p value < 0.1%), 0.496 (p value < 0.1%), 0.602 (p value < 0.1%), 0.664 (p value < 0.1%), and 0.290 (p value < 0.1%), respectively. The DARAs of "excessive working time" and "access to improved source of drinking water" remain stable temporally because their correlation coefficients are only -0.118 (p value < 1%) and -0.078 (p value < 5%). To sum up, in terms of the DARAs, social-topic relatedness is increasing, while environmental-topic relatedness is decreasing.

Table 2 lists the statistics of DANSA. News regarding "air pollution" and "access to improved source of drinking water" tends to be negative. **Figure 4** shows the temporal variation of DANSA. We estimate the correlations between the order of date and NSs of each topic. If the DANSA of a certain topic is significantly positively correlated with the order of date, it means that the media's attitude is gradually becoming positive on this topic, and vice versa. In terms of the correlation analysis, the DANSAs of "forced labor", "freedom of association, collective bargaining, right to strike", "unemployment", "poverty and inequality", "corruption", "excessive working time", and "access to health care" are gradually improving. Specifically, their correlation coefficients are 0.302 (p value < 0.1%), 0.293 (p value < 0.1%), 0.373 (p value < 0.1%), 0.223 (p value < 0.1%), 0.296 (p value < 0.1%), 0.401 (p value < 0.1%), and 0.222 (p value < 0.1%), respectively. The absolute values of the correlation coefficients of other topics are lower than 0.2, which are considered no significant correlations.

Table 3 summarizes the statistics of DARRs. The DARR is based on the news related to the topic. In other words, the NR towards a specific topic is required to be greater than 0. A higher mean of DARR towards a topic indicates that the potentially related news is highly stuck to the topic rather than simply mentioning some synonyms of the topic. Among 12 topics, the average DARR of "corruption" is the highest, 0.626. **Figure 5** demonstrates the temporal variation of DARRs. Except "air pollution", the DARRs of other topics are temporally increasing because their correlation coefficients of the topics listed in **Table 3** are 0.243 (p value < 0.1%), 0.332 (p value < 0.1%), 0.556 (p value < 0.1%), 0.456 (p value < 0.1%), 0.669 (p value < 0.1%), 0.647 (p value < 0.1%), 0.672 (p value < 0.1%), 0.334 (p value < 0.1%), 0.324 (p value < 0.1%), 0.477 (p value < 0.1%), and 0.329 (p value < 0.1%), respectively. **Figure 6** illustrates the daily new counts related to each topic. The DRNCs towards each topic are highly correlated with the daily news counts. The 12 correlation coefficients between daily news count and each topic are 0.838 (p value < 0.1%), 0.765 (p value < 0.1%), 0.727 (p value < 0.1%), 0.992 (p value < 0.1%), 0.894 (p value < 0.1%), 0.995 (p value < 0.1%), 0.849 (p value < 0.1%), 0.965 (p value < 0.1%), 0.681 (p value < 0.1%), 0.989 (p value < 0.1%), 0.871 (p value < 0.1%), and 0.971 (p value < 0.1%), respectively.

Table 4 notes the statistics of DANSR. Similar to the results in **Table 2**, on average, the NS of the news related to "air pollution" and "access to improved source of drinking water" tends to be negative, while others are apt to be positive. **Figure 7** expresses the temporal variations of the DANSRs towards each topic. The attitudes of "energy usage", "greenhouse gas", "force labor", "unemployment", "excessive working time", and "access to health care" are gradually becoming more positive because their correlation coefficients with the order of date are 0.406 (p value < 0.1%), 0.244 (p value < 0.1%), 0.316 (p value < 0.1%), 0.373 (p value < 0.1%), 0.436 (p value < 0.1%), and 0.203 (p value < 0.1%), respectively.

Discussion

This study is the first research to probe the temporal trends of ESG topics relatedness and attitude variations based on a large amount of available news. To investigate the correlations between news and a certain topic efficiently, objectively, and quantitatively, we create a powerful LLM, TMPT, according to the cutting-edge NLP technology, transformer, and the famous pre-trained model, BERT. In total, the TMPT model includes 138,843,049 parameters. Our model's accuracy is 85.73% in the cross-validation for the zero-shot labeling task. In other words, the TMPT could relatively accurately judge whether and how much the input news is related to the input topic, even though the model has never seen the topic during the training process. Our study investigates the NR and NS of 12 hot ESG-related topics based on 2,164,153 records of available news from more than 300 news sources in the U.S. from September 1st, 2021, to September 1st, 2023. According to data mining results, the media have been gradually publishing more news related to ESG topics since the DARAs and DARRs of most topics are increasing with time. Furthermore, the sentiment and attitude towards those topics are also becoming more and more positive, referring to temporal variations of the DANSAs and DANSRs.

The number of daily grasped news varies significantly. To answer whether this variation enormously impacts our analysis, we estimated the correlations between daily news count and the DRNCs of each topic. All DRNCs are highly related to the daily news count, which indicates the news is randomly and evenly missed. In this way, the impacts of the variation on this study could be ignored because we focus on the daily average values. We examine the correlations between DARAs and DANSAs and the daily news counts. The correlation coefficients are generally lower than 0.2, indicating that the missing news is randomly distributed.

According to all news in our database, news related to “air pollution” and “access to improved source of drinking water” usually show a negative tone, mirrored by the statistical indicators of their DANSAs. In addition, there is no trend in the content related to these two topics over time. During the period covered by our study, news mentions of these two topics were generally mostly critical or negative, and the media’s attention to them is stable over time. Since these two environmental issues are people’s basic survival needs, they have been attracting public attention with some fluctuations, consistent with previous studies on other countries (Lam, Cunsolo, Sawatzky, Ford, & Harper, 2017; Zhang et al., 2022). For the other ten topics, there are usually more reports with a positive tone than negative ones. Translated into reality, the coverage of the news related to these ten topics is mainly improvements and enhancements rather than criticism. Furthermore, the topics, “forced labor”, “freedom of association, collective bargaining, right to strike”, “unemployment”, “poverty and inequality”, “corruption”, “excessive working time”, and “access to health care” are becoming more positive over time. This phenomenon reflects that many issues, such as unemployment and excessive working, of public concern are being improved during this period. It is consistent with the society status: according to the data from the U.S. Bureau, from September 2021 to September 2023, the unemployment rate across society has continued to decline and gradually returned to pre-COVID-19 levels due to U.S. monetary policy (Bennani, 2023). Therefore, the whole society's views on employment have gradually become positive. Similarly, the information on the topics related to employment also gradually becomes cheerful.

News is a critical resource for investigating the public attitude towards a topic (Brandt & Gao, 2019; Dang, Moshirian, & Zhang, 2015). The news could not only reflect the market price of a certain industry but also lead to a change in the public attitude (Dang et al., 2020; Dang et al., 2019; Meng, Song, Liu, Wu, & Zeng, 2020). Previous studies claim that news sentiment could partially predict the economic indicators of a specific industry (Brandt & Gao, 2019; Rognone, Hyde, & Zhang, 2020). Sentiment analysis is definitely an ideal solution to analyze the news text. Previous studies mainly employ commercial databases, including labels and sentiment analysis results, to perform the analysis (Brandt & Gao, 2019; Dang et al., 2019). Beyond previous studies, we train a novel machine learning model, TMPT, to investigate the relatedness between the news and a selected topic. The database used by previous studies mainly relies on named entity recognition. Simply speaking, it judges whether the news is related to an entity. If it is, then the labels including location, industry, and others, would be conveyed to the news. In this way, in that database, they cannot examine the association of the news with a customized topic, at least not agilely. Deep learning gradually begins to be applied in news textual analysis, e.g., a study uses a fine-tuned BERT to label ESG topics based on a supervised learning strategy (Bang, Ryu, & Yu, 2023). Although the labeling process by the fine-tuned model is not artificial, the training dataset to train the model still requires human judgments. The zero-shot learning is free from humans even in the training dataset making. Furthermore, a zero-shot model is more flexible and could be used in other tasks.

Different from previous studies that mainly focus on one or several industries (Capelle-Blancard & Petit, 2019; Dang et al., 2020), we concentrate on specific topics of ESG. Of course, governments, society, and academia have been increasing voices on ESG. For example, the regulations, such as SA8000 (https://sa-intl.org/resources/sa8000-standard/) and guidelines of TNDF (https://tnfd.global/), are disclosed and executed (Birindelli & Chiappini, 2021). Moreover, in the past several years, ESG has been more gained more discussions in a variety of scopes in academia (Keeley et al., 2022; Senadheera et al., 2022). The detailed variations in topics of ESG could demonstrate more valuable information. Although the temperature of ESG is increasing, some topics are gradually less mentioned, namely environmental aspects. The COVID-19 pandemic has dramatically impacted the U.S. economy, exacerbating social problems (Alizadeh, Sharifi, Damanbagh, Nazarnia, & Nazarnia, 2023). Hence, the all news's relatedness to social topics is growing with time. Each media source has limited abilities to report. With an increase in the social news, other news would inevitably decrease. Furthermore, people are willing to pay more attention to solving environmental issues after reaching some incomes levels (Fujii, Iwata, Chapman, Kagawa, & Managi, 2018; C. Li & Managi, 2021; Wang, Yang, & Li, 2023). Our analysis proves that the change trends of the ESG topics might be inconsistent with the overall direction of ESG based on a vast dataset.

The results of this study offer substantial implications for academics, government policy-makers, company decision-makers, and investors, shedding light on the evolving landscape of ESG issues as reflected in media discourse. For academics, the findings provide a solid quantitative foundation for understanding media coverage and public perceptions of ESG topics and how they change over time. This could inspire further research into the mechanisms through which media narratives shape public and corporate ESG agendas. Additionally, our research provides a technical paradigm for extensive and efficient big data text analysis via LLMs. Through a large amount of news and LLM based on our process, scholars can not only track ESG-related topics but also analyze other topics to provide more perspectives. For policy-making and corporate strategy, these insights highlight the importance of media monitoring to gauge public sentiment and adjust strategies accordingly (Baier et al., 2020; Rognone et al., 2020). Companies can leverage this knowledge to better align their ESG initiatives with public concerns, enhancing their reputational capital and ensuring compliance with societal expectations (Keeley et al., 2022). Furthermore, by understanding the media's role in shaping ESG perceptions, companies can more effectively communicate their ESG commitments, fostering greater trust and engagement with stakeholders.

There are several limitations of this study should be noted. First, we directly employ a publicly available model from HuggingFace without further fine-tuning. Although we select the best model in terms of the scope, task, performance, and structure, it is still far from perfect. The current model is trained by a 20-thousand-record dataset of the financial news. The current model's input is required no more than 512 tokens, or roughly 400 words. Second, the TMPT's performance is approximately 85%, which is impressive in zero-shot learning. By increasing the depth of the networks and using more data to train, the TMPT has space to improve. Third, we miss a batch of news data, due to the grasping ability of the current data vendor. Fourth, currently, we only focus on the English language news owing to our training logic, training dataset, and the embedding layer from BERT. In further studies, a new model of news sentiment analysis would be trained. A multi-language and more accurate TMPT is also considered and planned, which could cover global analysis effectively and efficiently.

Conclusions

According to the analysis of 2,164,153 new records published in the U.S. from more than 300 news sources from September 1st, 2021, to September 1st, 2023, the media gradually emphasized social issues, while the voice on environmental problems became lower and lower. These change trends mainly express the relatedness between the news and the selected topics based on the daily summary. In other words, for some topics, media coverage is increasingly focused on clear topics. Moreover, the attitudes towards each topic of interest are incrementally becoming cheerful. To quantitively, objectively, and effectively probe the public's attitude towards ESG topics, we build a powerful machine learning model, TMPT. This study demonstrates the temporal variation in the public attitude towards 12 ESG topics, which offer critical information to governments, businesses, and societies.

Acknowledgments

This research was supported by the following funding agencies: JSPS KAKENHI (Grant No. JP20H00648), the Environment Research and Technology Development Fund of the Environmental Restoration and Conservation Agency of Japan (Grant No. JPMEERF20201001), and JST MiraiProgram (Grant No. JPMJMI22I4).

Figure

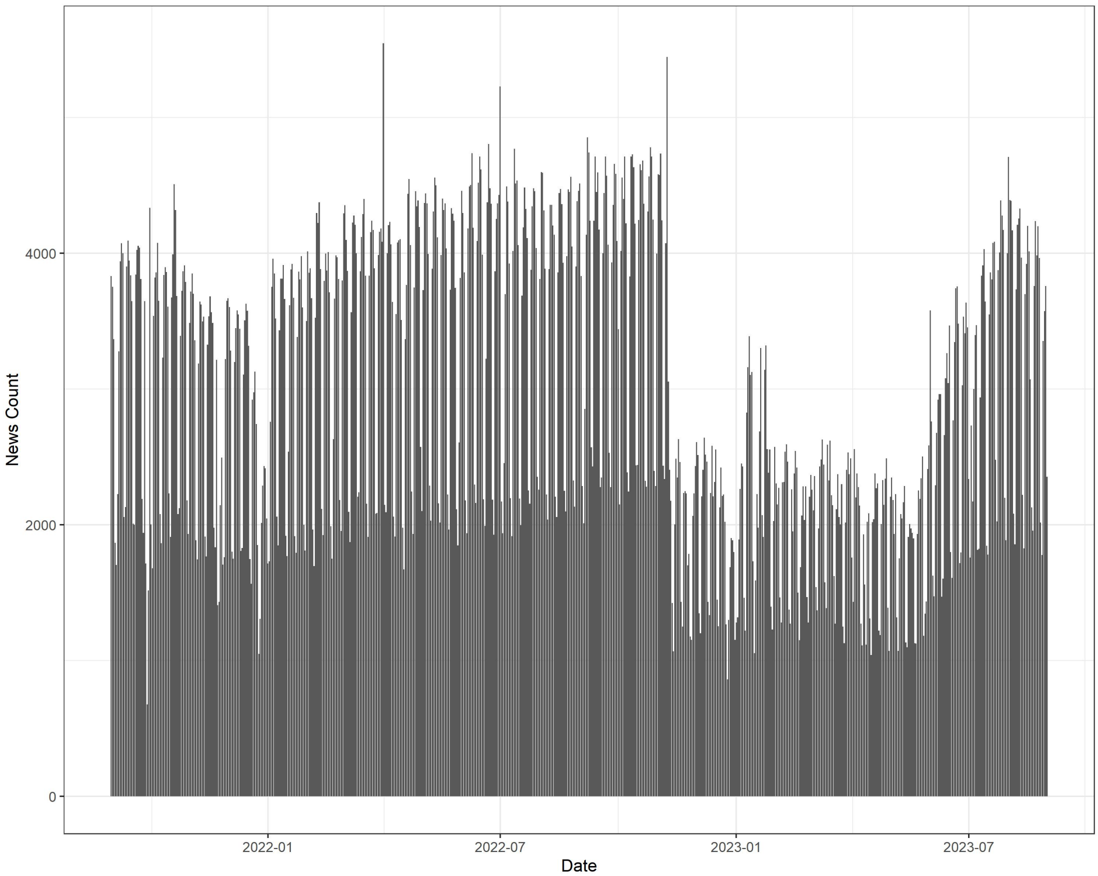

Figure 1: Daily News Count

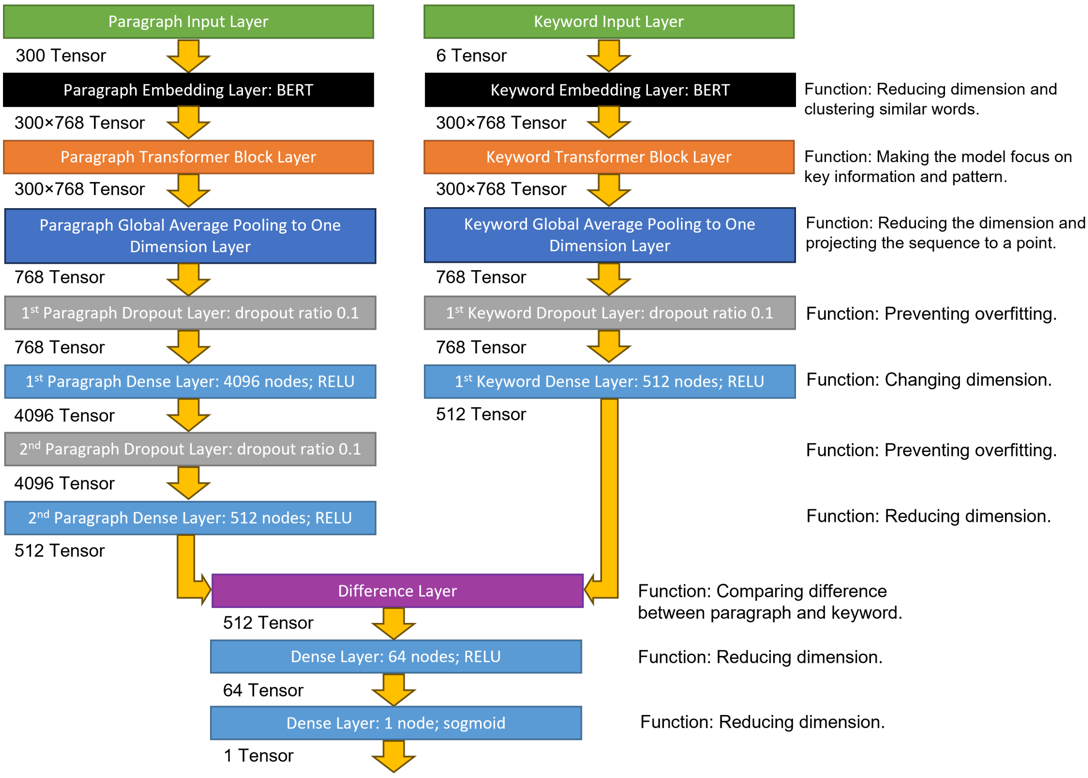

Figure 2: Architecture of the TMPT Model

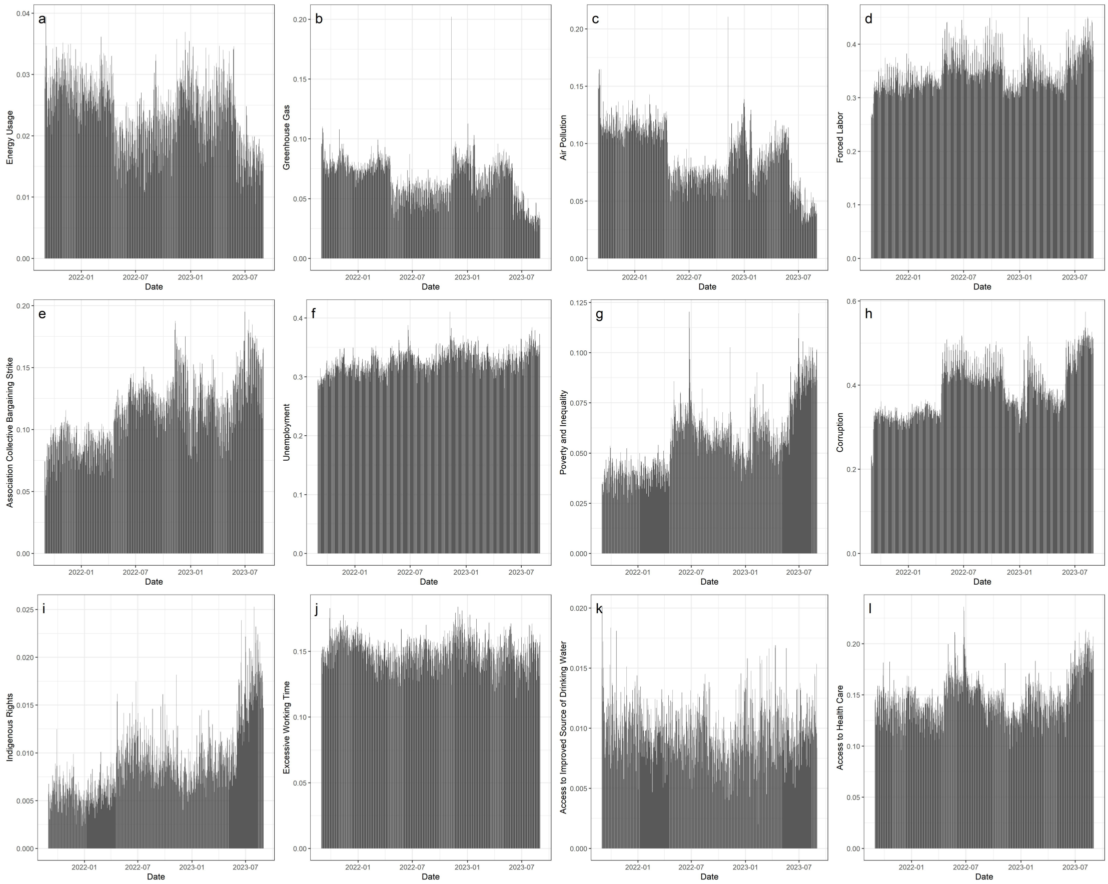

Figure 3: Temporal Variations of DARAs

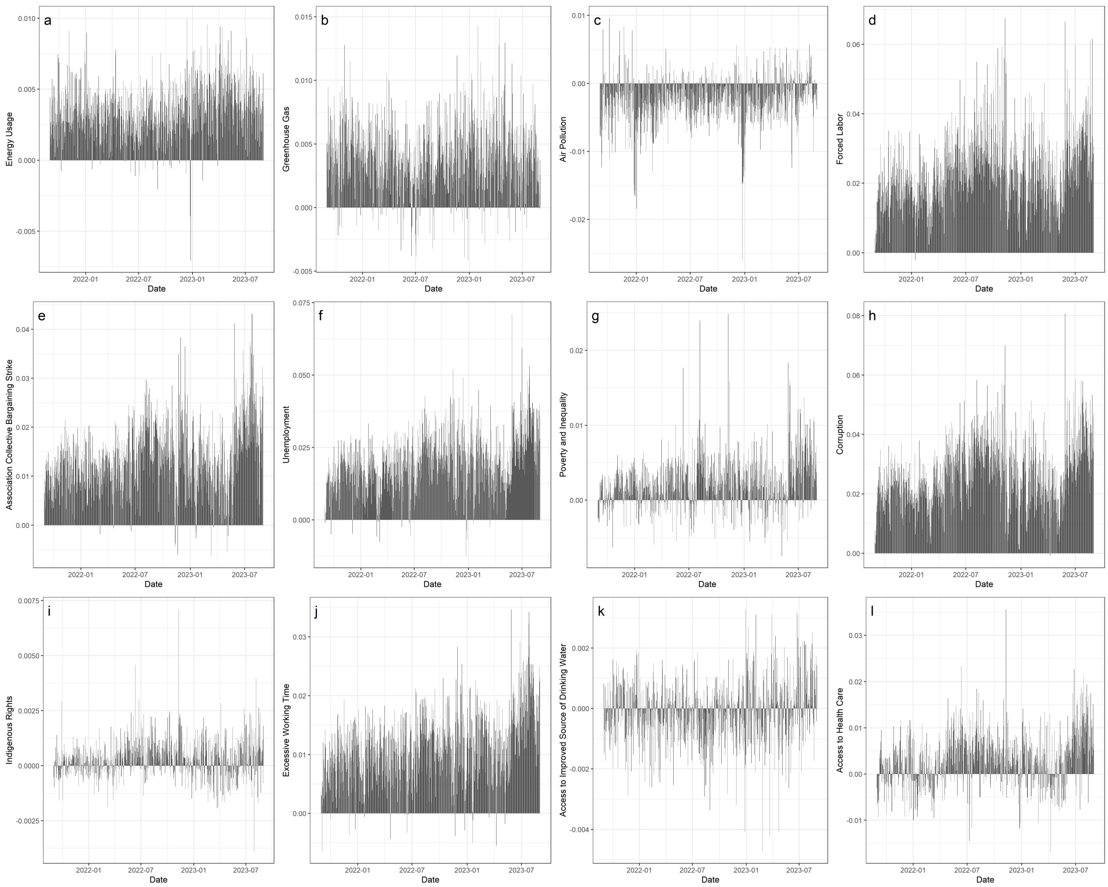

Figure 4: Temporal Variations of DANSAs

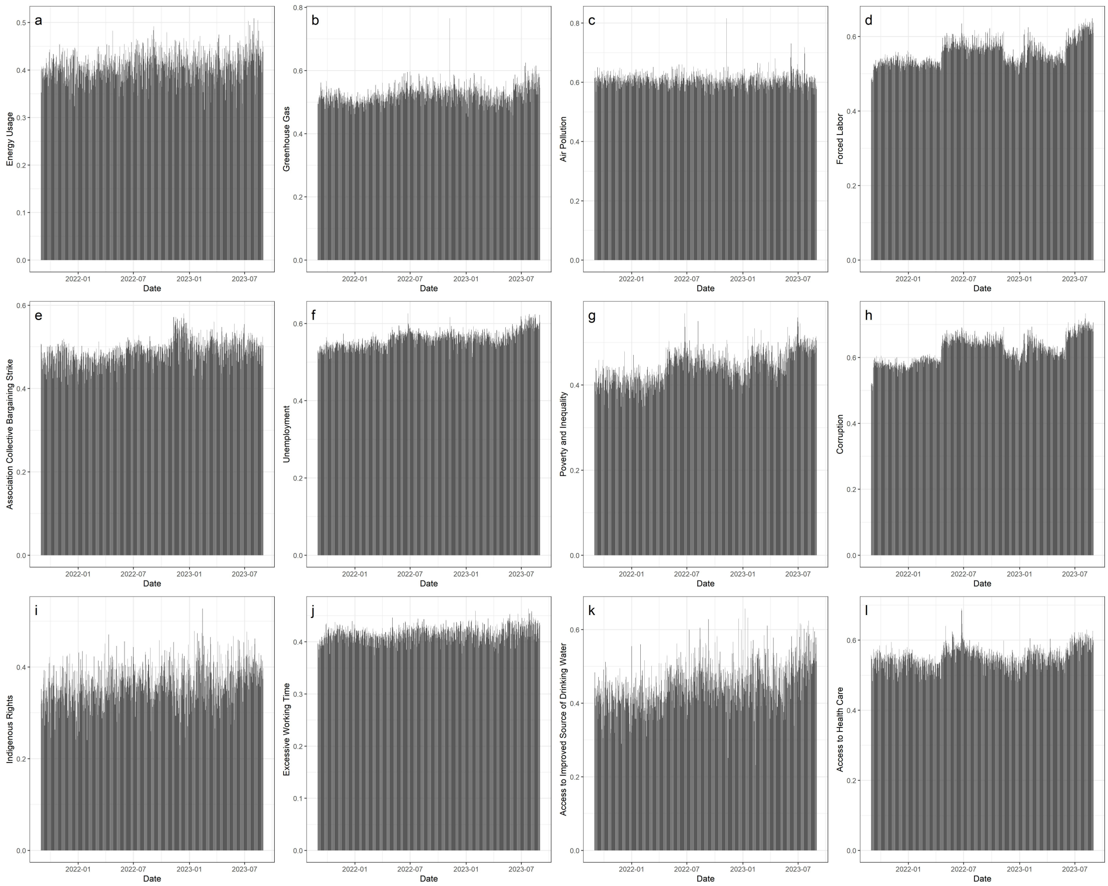

Figure 5: Temporal Variations of DARRs

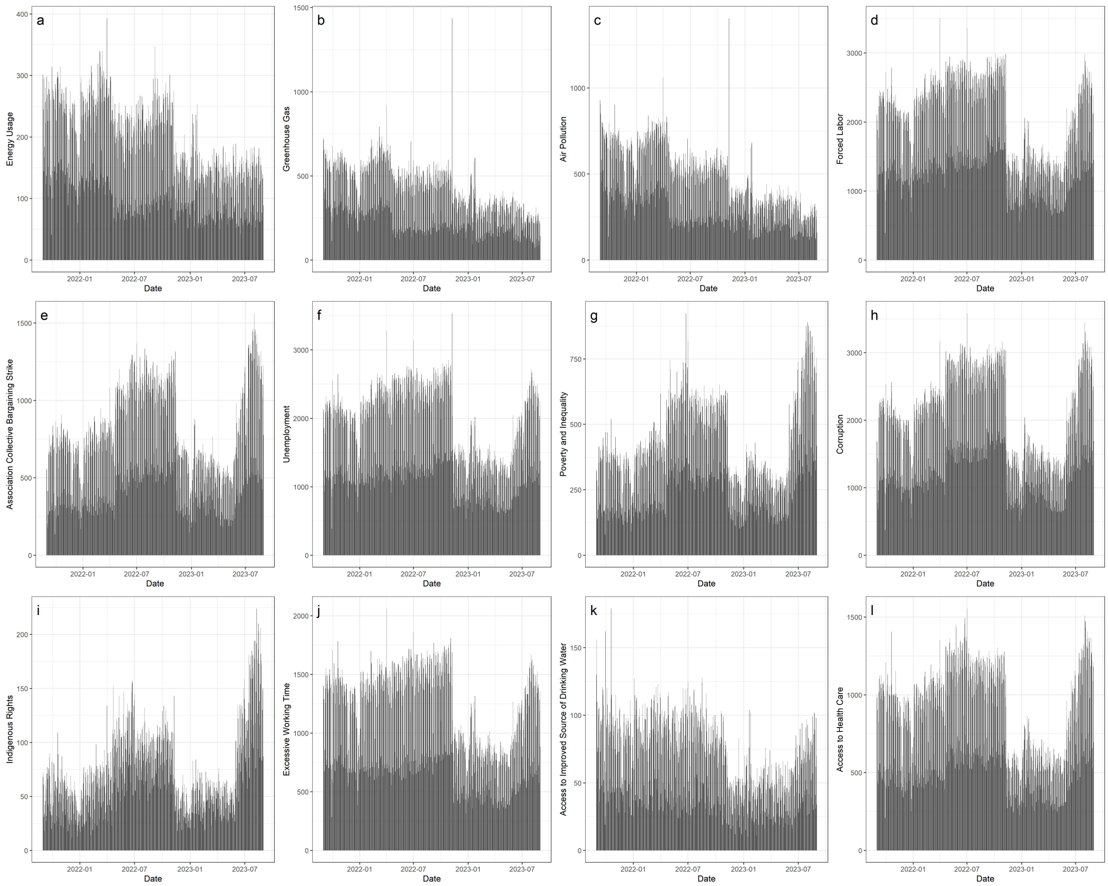

Figure 6: DRNCs towards Each Topic

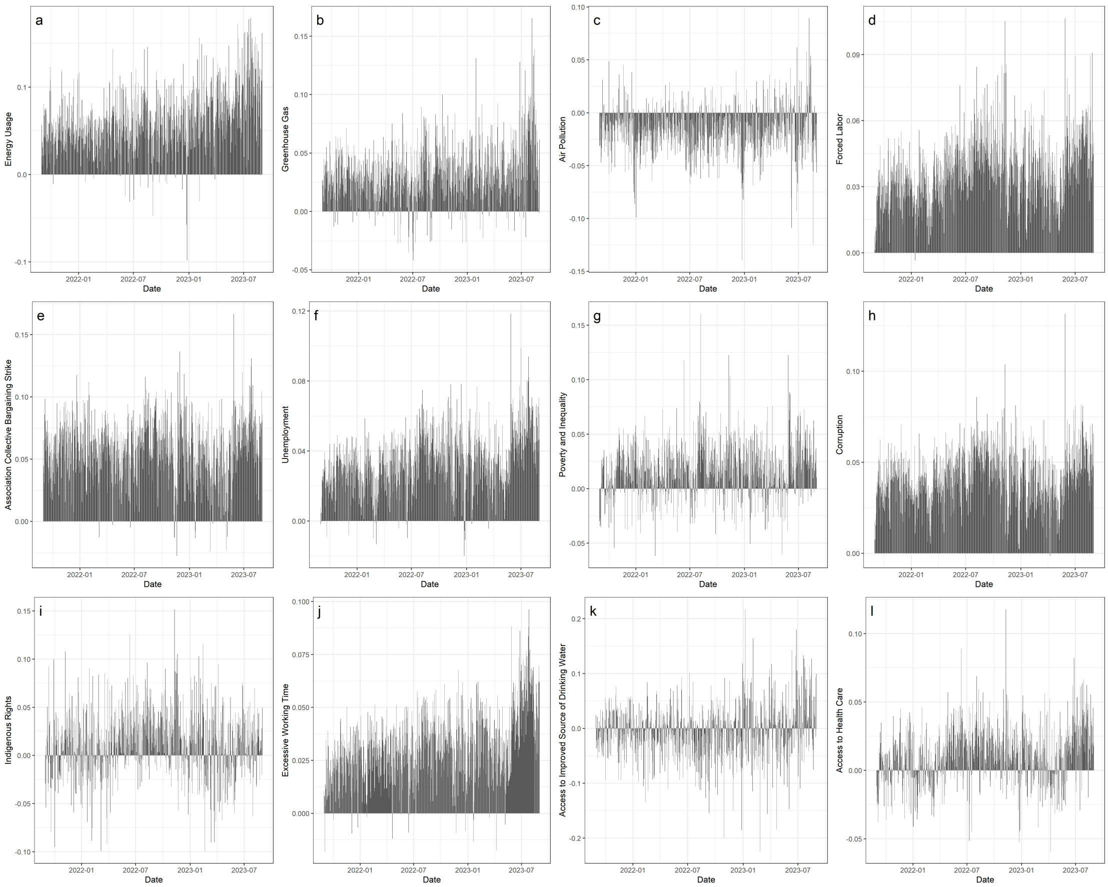

Figure 7: Temporal Variations of DANSRs

Table

Table 1: Summary of DARA Statistics

Statistic	N	Mean	St. Dev.	Min	Median	Max
Energy usage	730	0.024	0.006	0.009	0.024	0.039
Greenhouse gas	730	0.067	0.018	0.022	0.070	0.202
Air pollution	730	0.090	0.027	0.029	0.090	0.211
Forced labor	730	0.349	0.034	0.261	0.343	0.451
Freedom of association, collective bargaining, right to strike	730	0.115	0.027	0.047	0.116	0.195
Unemployment	730	0.328	0.019	0.279	0.328	0.411
Poverty and inequality	730	0.057	0.017	0.025	0.056	0.120
Corruption	730	0.392	0.064	0.207	0.384	0.574
Indigenous rights	730	0.009	0.004	0.002	0.008	0.025
Excessive working time	730	0.151	0.011	0.115	0.152	0.184
Access to improved source of drinking water	730	0.010	0.002	0.002	0.010	0.020
Access to health care	730	0.148	0.021	0.096	0.147	0.236

Table 2: Summary of DANSA Statistics

Statistic	N	Mean	St. Dev.	Min	Median	Max
NS of Energy usage	730	0.004	0.002	-0.007	0.003	0.010
NS of Greenhouse gas	730	0.004	0.003	-0.004	0.004	0.015
NS of Air pollution	730	-0.003	0.004	-0.026	-0.003	0.010
NS of Forced labor	730	0.023	0.010	-0.002	0.023	0.067
NS of Freedom of association, collective bargaining, right to strike	730	0.014	0.007	-0.006	0.013	0.043
NS of Unemployment	730	0.020	0.010	-0.012	0.020	0.071
NS of Poverty and inequality	730	0.002	0.004	-0.007	0.002	0.025
NS of Corruption	730	0.027	0.011	-0.001	0.026	0.081
NS of Indigenous rights	730	0.0003	0.001	-0.004	0.0003	0.007
NS of Excessive working time	730	0.011	0.006	-0.006	0.011	0.035
NS of Access to improved source of drinking water	730	-0.0001	0.001	-0.005	-0.0001	0.003
NS of Access to health care	730	0.003	0.006	-0.017	0.003	0.036

Table 3: Summary of DARR Statistics

Statistic	N	Mean	St. Dev.	Min	Median	Max
Energy usage	730	0.413	0.029	0.316	0.412	0.509
Greenhouse gas	730	0.527	0.029	0.454	0.525	0.766
Air pollution	730	0.609	0.022	0.540	0.609	0.816
Forced labor	730	0.559	0.033	0.478	0.552	0.649
Freedom of association, collective bargaining, right to strike	730	0.491	0.029	0.409	0.492	0.580
Unemployment	730	0.562	0.022	0.507	0.562	0.626
Poverty and inequality	730	0.447	0.037	0.345	0.446	0.568
Corruption	730	0.626	0.041	0.499	0.628	0.734
Indigenous rights	730	0.369	0.042	0.229	0.367	0.527
Excessive working time	730	0.417	0.014	0.370	0.418	0.464
Access to improved source of drinking water	730	0.457	0.060	0.232	0.453	0.657
Access to health care	730	0.554	0.030	0.480	0.553	0.690

Table 4: Summary of DANSR Statistics

Statistic	N	Mean	St. Dev.	Min	Median	Max
NS of Energy usage	730	0.063	0.036	-0.098	0.061	0.179
NS of Greenhouse gas	730	0.034	0.025	-0.042	0.032	0.165
NS of Air pollution	730	-0.018	0.024	-0.140	-0.018	0.089
NS of Forced labor	730	0.037	0.016	-0.003	0.037	0.106
NS of Freedom of association, collective bargaining, right to strike	730	0.059	0.025	-0.028	0.061	0.167
NS of Unemployment	730	0.035	0.017	-0.020	0.034	0.118
NS of Poverty and inequality	730	0.018	0.026	-0.062	0.017	0.161
NS of Corruption	730	0.043	0.015	-0.001	0.042	0.132
NS of Indigenous rights	730	0.012	0.035	-0.100	0.013	0.152
NS of Excessive working time	730	0.032	0.017	-0.018	0.031	0.096
NS of Access to improved source of drinking water	730	-0.007	0.052	-0.224	-0.006	0.217
NS of Access to health care	730	0.011	0.022	-0.059	0.012	0.118

Reference:

Aburto, J. M., Tilstra, A. M., Floridi, G., & Dowd, J. B. (2022). Significant impacts of the COVID-19 pandemic on race/ethnic differences in US mortality. *Proceedings of the National Academy of Sciences, 119*(35), e2205813119. doi:10.1073/pnas.2205813119

Alizadeh, H., Sharifi, A., Damanbagh, S., Nazarnia, H., & Nazarnia, M. (2023). Impacts of the COVID-19 pandemic on the social sphere and lessons for crisis management: a literature review. *Natural Hazards, 117*(3), 2139-2164. doi:10.1007/s11069-023-05959-2

Anita, M., Shveta, S., Yadav Surendra, S., & Arvind, M. (2023). When do ESG controversies reduce firm value in India? *Global Finance Journal, 55*, 100809. doi:https://doi.org/10.1016/j.gfj.2023.100809

Baier, P., Berninger, M., & Kiesel, F. (2020). Environmental, social and governance reporting in annual reports: A textual analysis. *Financial Markets, Institutions & Instruments, 29*(3), 93-118. doi:10.1111/fmii.12132

Bang, J., Ryu, D., & Yu, J. (2023). ESG controversies and investor trading behavior in the Korean market. *Finance Research Letters, 54*, 103750. doi:https://doi.org/10.1016/j.frl.2023.103750

Bennani, H. (2023). Effect of monetary policy shocks on the racial unemployment rates in the US. *Economic Systems, 47*(1), 101058. doi:https://doi.org/10.1016/j.ecosys.2022.101058

Birindelli, G., & Chiappini, H. (2021). Climate change policies: Good news or bad news for firms in the European Union? *Corporate Social Responsibility and Environmental Management, 28*(2), 831-848. doi:10.1002/csr.2093

Brandt, M. W., & Gao, L. (2019). Macro fundamentals or geopolitical events? A textual analysis of news events for crude oil. *Journal of Empirical Finance, 51*, 64-94. doi:https://doi.org/10.1016/j.jempfin.2019.01.007

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., . . . Askell, A. (2020). Language models are few-shot learners. *Advances in neural information processing systems, 33*, 1877-1901.

Capelle-Blancard, G., & Petit, A. (2019). Every Little Helps? ESG News and Stock Market Reaction. *Journal of Business Ethics, 157*(2), 543-565. doi:10.1007/s10551-017-3667-3

Cezarino, L. O., Liboni, L. B., Hunter, T., Pacheco, L. M., & Martins, F. P. (2022). Corporate social responsibility in emerging markets: Opportunities and challenges for sustainability integration. *Journal of Cleaner Production, 362*, 132224. doi:https://doi.org/10.1016/j.jclepro.2022.132224

Cohen, G. (2023). The impact of ESG risks on corporate value. *Review of Quantitative Finance and Accounting, 60*(4), 1451-1468. doi:10.1007/s11156-023-01135-6

Dang, T. L., Dang, M., Hoang, L., Nguyen, L., & Phan, H. L. (2020). Media coverage and stock price synchronicity. *International Review of Financial Analysis, 67*, 101430. doi:https://doi.org/10.1016/j.irfa.2019.101430

Dang, T. L., Dang, V. A., Moshirian, F., Nguyen, L., & Zhang, B. (2019). News media coverage and corporate leverage adjustments. *Journal of Banking & Finance, 109*, 105666. doi:https://doi.org/10.1016/j.jbankfin.2019.105666

Dang, T. L., Moshirian, F., & Zhang, B. (2015). Commonality in news around the world. *Journal of Financial Economics, 116*(1), 82-110. doi:https://doi.org/10.1016/j.jfineco.2014.11.007

De Giuli, M. E., Grechi, D., & Tanda, A. (2023). What do we know about <scp>ESG</scp> and risk? A systematic and bibliometric review. *Corporate Social Responsibility and Environmental Management*. doi:10.1002/csr.2624

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Ding, X., Appolloni, A., & Shahzad, M. (2022). Environmental administrative penalty, corporate environmental disclosures and the cost of debt. *Journal of Cleaner Production, 332*, 129919. doi:https://doi.org/10.1016/j.jclepro.2021.129919

Douthit, N., Kiv, S., Dwolatzky, T., & Biswas, S. (2015). Exposing some important barriers to health care access in the rural USA. *Public Health, 129*(6), 611-620. doi:https://doi.org/10.1016/j.puhe.2015.04.001

Fiaschi, D., Giuliani, E., Nieri, F., & Salvati, N. (2020). How bad is your company? Measuring corporate wrongdoing beyond the magic of ESG metrics. *Business Horizons, 63*(3), 287-299. doi:https://doi.org/10.1016/j.bushor.2019.09.004

Fujii, H., Iwata, K., Chapman, A., Kagawa, S., & Managi, S. (2018). An analysis of urban environmental Kuznets curve of CO2 emissions: Empirical analysis of 276 global metropolitan areas. *Applied Energy, 228*, 1561-1568. doi:10.1016/j.apenergy.2018.06.158

Geng, T., Jia, F., Cai, W., Wu, L., Gan, B., Jing, Z., . . . McPhaden, M. J. (2023). Increased occurrences of consecutive La Niña events under global warming. *Nature, 619*(7971), 774-781. doi:10.1038/s41586-023-06236-9

Gervais, E., Kleijn, R., Nold, S., & van der Voet, E. (2023). Risk-based due diligence in supply chains: The case of silver for photovoltaics. *Resources, Conservation and Recycling, 198*, 107148. doi:https://doi.org/10.1016/j.resconrec.2023.107148

Gibson Brandon, R., Glossner, S., Krueger, P., Matos, P., & Steffen, T. (2022). Do Responsible Investors Invest Responsibly?*. *Review of Finance, 26*(6), 1389-1432. doi:10.1093/rof/rfac064

Gregory, R. P. (2022). ESG scores and the response of the S&P 1500 to monetary and fiscal policy during the Covid-19 pandemic. *International Review of Economics & Finance, 78*, 446-456. doi:https://doi.org/10.1016/j.iref.2021.12.013

He, X., Jing, Q., & Chen, H. (2023). The impact of environmental tax laws on heavy-polluting enterprise ESG performance: A stakeholder behavior perspective. *Journal of Environmental Management, 344*, 118578. doi:https://doi.org/10.1016/j.jenvman.2023.118578

Hurlbert, M. (2020). Access and allocation: rights to water, sanitation and hygiene. *International Environmental Agreements: Politics, Law and Economics, 20*(2), 339-358. doi:10.1007/s10784-020-09484-6

Keeley, A. R., Chapman, A. J., Yoshida, K., Xie, J., Imbulana, J., Takeda, S., & Managi, S. (2022). ESG metrics and social equity: Investigating commensurability. *Frontiers in Sustainability, 3*, 920955.

Lam, S., Cunsolo, A., Sawatzky, A., Ford, J., & Harper, S. L. (2017). How does the media portray drinking water security in Indigenous communities in Canada? An analysis of Canadian newspaper coverage from 2000-2015. *BMC Public Health, 17*(1), 282. doi:10.1186/s12889-017-4164-4

Le Billon, P., Lujala, P., Singh, D., Culbert, V., & Kristoffersen, B. (2021). Fossil fuels, climate change, and the COVID-19 crisis: pathways for a just and green post-pandemic recovery. *Climate Policy, 21*(10), 1347-1356.

Le, T. T., Vo, X. V., & Venkatesh, V. G. (2022). Role of green innovation and supply chain management in driving sustainable corporate performance. *Journal of Cleaner Production, 374*, 133875. doi:https://doi.org/10.1016/j.jclepro.2022.133875

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., & Pozzer, A. (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature, 525*(7569), 367-371. doi:10.1038/nature15371

Lerouge, R., Lema, M. D., & Arnaboldi, M. (2023). The role played by government communication on the level of public fear in social media: An investigation into the Covid-19 crisis in Italy. *Government Information Quarterly, 40*(2), 101798. doi:https://doi.org/10.1016/j.giq.2022.101798

Li, C., & Managi, S. (2021). Spatial Variability of the Relationship between Air Pollution and Well-being. *Sustainable Cities and Society*, 103447. doi:10.1016/j.scs.2021.103447

Li, X., Huang, X., Li, D., & Xu, Y. (2022). Aggravated social segregation during the COVID-19 pandemic: Evidence from crowdsourced mobility data in twelve most populated U.S. metropolitan areas. *Sustainable Cities and Society, 81*, 103869. doi:https://doi.org/10.1016/j.scs.2022.103869

Litvinenko, V., Bowbrick, I., Naumov, I., & Zaitseva, Z. (2022). Global guidelines and requirements for professional competencies of natural resource extraction engineers: Implications for ESG principles and sustainable development goals. *Journal of Cleaner Production, 338*, 130530. doi:https://doi.org/10.1016/j.jclepro.2022.130530

Liu, M., Luo, X., & Lu, W.-Z. (2023). Public perceptions of environmental, social, and governance (ESG) based on social media data: Evidence from China. *Journal of Cleaner Production, 387*, 135840. doi:https://doi.org/10.1016/j.jclepro.2022.135840

Luu, J., & Palczewski, K. (2018). Human aging and disease: Lessons from age-related macular degeneration. *Proceedings of the National Academy of Sciences, 115*(12), 2866-2872. doi:10.1073/pnas.1721033115

Malo, P., Sinha, A., Korhonen, P., Wallenius, J., & Takala, P. (2014). Good debt or bad debt: Detecting semantic orientations in economic texts. *Journal of the Association for Information Science and Technology, 65*(4), 782-796.

Marrero, M., Urbano, J., Sánchez-Cuadrado, S., Morato, J., & Gómez-Berbís, J. M. (2013). Named Entity Recognition: Fallacies, challenges and opportunities. *Computer Standards & Interfaces, 35*(5), 482-489. doi:https://doi.org/10.1016/j.csi.2012.09.004

Meng, Q., Song, X., Liu, C., Wu, Q., & Zeng, H. (2020). The impact of block trades on stock price synchronicity: Evidence from China. *International Review of Economics & Finance, 68*, 239-253. doi:https://doi.org/10.1016/j.iref.2020.04.009

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems, 26*.

Moore, R. C., Dahlke, R., & Hancock, J. T. (2023). Exposure to untrustworthy websites in the 2020 US election. *Nature Human Behaviour, 7*(7), 1096-1105. doi:10.1038/s41562-023-01564-2

Ng, T.-H., Lye, C.-T., Chan, K.-H., Lim, Y.-Z., & Lim, Y.-S. (2020). Sustainability in Asia: The Roles of Financial Development in Environmental, Social and

Governance (ESG) Performance. *Social Indicators Research, 150*(1), 17-44. doi:10.1007/s11205-020-02288-w
Nobanee, H., & Ellili, N. (2016). Corporate sustainability disclosure in annual reports: Evidence from UAE banks: Islamic versus conventional. *Renewable and Sustainable Energy Reviews, 55*, 1336-1341. doi:https://doi.org/10.1016/j.rser.2015.07.084
Pan, C., Abbas, J., Álvarez-Otero, S., Khan, H., & Cai, C. (2022). Interplay between corporate social responsibility and organizational green culture and their role in employees' responsible behavior towards the environment and society. *Journal of Cleaner Production, 366*, 132878. doi:https://doi.org/10.1016/j.jclepro.2022.132878
Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog, 1*(8), 9.
Rognone, L., Hyde, S., & Zhang, S. S. (2020). News sentiment in the cryptocurrency market: An empirical comparison with Forex. *International Review of Financial Analysis, 69*, 101462. doi:https://doi.org/10.1016/j.irfa.2020.101462
Senadheera, S. S., Gregory, R., Rinklebe, J., Farrukh, M., Rhee, J. H., & Ok, Y. S. (2022). The development of research on environmental, social, and governance (ESG): A bibliometric analysis. *Sustainable Environment, 8*(1). doi:10.1080/27658511.2022.2125869
Ullah, S., Adams, K., Adams, D., & Attah-Boakye, R. (2021). Multinational corporations and human rights violations in emerging economies: Does commitment to social and environmental responsibility matter? *Journal of Environmental Management, 280*, 111689. doi:https://doi.org/10.1016/j.jenvman.2020.111689
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Aidan, . . . Polosukhin, I. (2017). Attention Is All You Need. *arXiv pre-print server*. doi:None

arxiv:1706.03762
Vo, T. P. T., Ngo, H. H., Guo, W., Turney, C., Liu, Y., Nguyen, D. D., . . . Varjani, S. (2023). Influence of the COVID-19 pandemic on climate change summit negotiations from the climate governance perspective. *Science of The Total Environment, 878*, 162936. doi:https://doi.org/10.1016/j.scitotenv.2023.162936
Wang, Q., Yang, T., & Li, R. (2023). Does income inequality reshape the environmental Kuznets curve (EKC) hypothesis? A nonlinear panel data analysis. *Environmental Research, 216*, 114575. doi:https://doi.org/10.1016/j.envres.2022.114575
Wekhof, T., & Houde, S. (2023). Using narratives to infer preferences in understanding the energy efficiency gap. *Nature Energy*. doi:10.1038/s41560-023-01303-x
Weston, P., & Nnadi, M. (2023). Evaluation of strategic and financial variables of corporate sustainability and ESG policies on corporate finance performance. *Journal of Sustainable Finance & Investment, 13*(2), 1058-1074.
WHO. (2020). *WHO global water, sanitation and hygiene: annual report 2019* (9240013393). Retrieved from
York, A. M., Zdor, E., BurnSilver, S., Degai, T., Monakhova, M., Isakova, S., . . . Kempf, M. (2022). Institutional navigation of oceans governance: Lessons from Russia and the United States Indigenous multi-level whaling governance in the Arctic. *Earth System Governance, 14*, 100154. doi:https://doi.org/10.1016/j.esg.2022.100154

Yu, H., Liang, C., Liu, Z., & Wang, H. (2023). News-based ESG sentiment and stock price crash risk. *International Review of Financial Analysis, 88*, 102646. doi:https://doi.org/10.1016/j.irfa.2023.102646

Zhang, D., Pan, S. L., Yu, J., & Liu, W. (2022). Orchestrating big data analytics capability for sustainability: A study of air pollution management in China. *Information & Management, 59*(5), 103231. doi:https://doi.org/10.1016/j.im.2019.103231