Junior Management Science 10(2) (2025) 369-401

Junior Management Science

www.jums.academy
ISSN: 2942-1861

# Sustainability in the Corporate Sector: A News Textual Analysis Approach to Measuring ESG Performance

Mohammad Izzat Raihan Imron

*Technical University of Munich*

**Abstract**

Sustainability has become a crucial factor in the financial sector, making the assessment of a company's sustainability performance essential for informed decision-making. Recognizing the media's power to shape public perception of corporate sustainability issues, this study examines the use of news analysis to evaluate companies' performance against Environmental, Social, and Governance (ESG) criteria. Leveraging OpenAI's models, this research parses unstructured data within news articles and introduces a machine learning pipeline to score companies' ESG performance based on their media representation. The study uncovers several key findings: firstly, it demonstrates that a less costly, fine-tuned model can surpass the zero-shot capabilities of a more expensive model in classifying ESG content. Secondly, it identifies discrepancies in media coverage across industries, leading to unequal assessments of companies. Thirdly, it reveals a media tendency to underreport companies' environmental efforts. Finally, the study highlights areas where companies face media criticism, suggesting potential improvements in their ESG practices. These insights contribute to the understanding of how machine learning can assist in the critical evaluation of sustainability in the business domain.

*Keywords:* ESG; machine learning; natural language processing; news; NLP; sustainability

## 1. Introduction

Over the past decade, sustainability has become a prominent topic in the corporate finance world. The burgeoning interest in integrating values into the investment decision-making process underscores its significance. Although there is still a lack of clarity about the meaning, and it has been referred to by various names (Starks, 2023), there has been a remarkable shift in investors' focus from short-term profits to the consideration of long-term impacts and the non-financial risks associated with the ventures (Krappel et al., 2021).

Sustainability factors, including environmental stewardship, co-prosperity, ethical responsibility, and the preservation of a company's economic performance, are gaining increasing significance (J. Lee & Kim, 2023). As a result, environmental, social, and governance (ESG) practices have surfaced as a novel corporate management paradigm. The ESG concept was first introduced in 'Who Cares Wins,' published by the United Nations Global Compact in 2004, where it integrated the dimensions of ESG and the Principles for Responsible Investment (PRI) made the concepts become more popular in the following years (J. Lee & Kim, 2023; United Nations Global Compact, 2004).

According to PwC's Asset and Wealth Management Revolution 2022 Report, ESG-related assets under management (AuM) were at US$18.4 trillion in 2021 and are expected to reach US$33.9 trillion in 2026, making up 21.5% of assets under management (PwC, 2022). This substantial growth is more than a financial trend. It reflects a meaningful change in investor attitudes towards sustainable investing. Hartzmark and Sussman (2019) affirm this perspective, and they argue that sustainability is a valued attribute among investors, influencing their investment behaviors. Investors are navigating their decisions not solely based on financial returns but are also influenced by emotional and ethical factors that resonate with their moral values (Hartzmark & Sussman, 2019). Empirical evidence indicates that investors react more positively than usual to the latest positive news

DOI: https://doi.org/10.5282/jums/v10i2pp369-401

for a company with superior sustainability performance, subsequently elevating the company's stock price (Barka et al., 2023; Leite & Uysal, 2023). Consequently, a firm's commitment to sustainability strengthens its reputation while simultaneously enhancing investors' perceptions.

Assessing a company's sustainability performance has become an integral aspect of investment decision-making, intertwining financial prospects with value considerations. A collective of more than five thousand investors, managing a cumulative of over US$100 trillion in assets, have pledged to incorporate ESG data into their investment decision-making processes (Principles for Responsible Investment (PRI), 2023). Investors often resort to ESG scores or ratings provided by various institutions to measure a firm's sustainability quality. Organizations like MSCI, Sustainalytics, and Refinitiv offer ESG ratings that evaluate companies based on various sustainability metrics. However, it is critical to note that ESG ratings, despite their growing usage and influence, are not a perfect method for determining a company's commitment to or effectiveness in sustainability efforts.

A number of critiques emphasize that ESG ratings come with their own specific challenges and shortcomings. Ilango (2023), for example, argues that the system has the potential to disrupt financial market stability if it remains unregulated. A Financial Times article by Allen (2018) underscores that ESG ratings might be overly simplified and subject to subjective measurement. Chatterji et al. (2016) show a notable inconsistency in social ratings provided by six renowned raters, with disparities persisting even after adjusting for variances in the definition of Corporate Social Responsibility (CSR) among them, suggesting that the ratings might have low validity. Berg et al. (2022) describe further that the cause of the discrepancy is not just different definitions but a fundamental disagreement regarding the underlying data. The issue of inconsistency is prevalent not only between the raters but also within a single rater across different years. Results from Berg et al. (2020) reveal that there have been considerable adjustments to the historical evaluations by Refinitiv ESG (formerly ASSET4).

ESG rating providers, in an effort to be different, are likely to continue using their own methodologies and metrics, intentionally creating discrepancies in ratings (Brackley et al., 2022). Coupled with a lack of transparency in the disclosure of rating criteria and calculation methods, raises concerns about the reliable use of ESG ratings (J. Lee & Kim, 2023). It is important to directly gather ESG-related information from various data sources to effectively obtain objective information on a company's ESG efforts (J. Lee & Kim, 2023). However, collecting unstructured, scattered, and vast information can be arduous and financially demanding. To alleviate this, Natural Language Processing (NLP) can serve as a supportive tool to help streamline this process.

Several studies have explored the use of Natural Language Processing (NLP) to extract information from text, particularly in financial and environmental contexts. In the broad financial context, A. H. Huang et al. (2023) introduce FinBERT, a language model tailored for finance-related text. In a more niche context, Webersinke et al. (2021) develop ClimateBERT, a model specializing in climate-related texts. In the ESG realm, various recent studies such as those by H. Kang and Kim (2022), J. Lee and Kim (2023), Luccioni et al. (2020), Mehra et al. (2022), and Polignano et al. (2022) have utilized diverse NLP methods to analyze companies' sustainability reports, highlighting a growing interest in leveraging this tool to comprehend and evaluate ESG performances.

This study introduces a different approach to understanding companies' ESG performance by employing a third-person perspective, particularly through news sources. A modest amount of research has explored applying NLP to news text in the context of ESG. For instance, Nugent et al. (2020) introduce $\text{BERT}_{RNA}$, which is capable of performing multi-class ESG controversy classification tasks. Additionally, Fischbach et al. (2022) develop ESG-Miner, based on the BERT model, designed to classify the headlines of news articles from Twitter in the ESG context. A more detailed description of these studies is explained in Section 2.

According to the agenda-setting theory, the news media plays a crucial role in directing public attention to significant issues and shaping perceptions and knowledge about those topics, influencing how the public views and understands them (McCombs & Reynolds, 2002). It is also true in the ESG context that mass media, including news, has the ability to drive people's perceptions regarding the firms' environmental performance and other ESG issues (N. Brown & Deegan, 1998; Hammami & Hendijani Zadeh, 2019). The results of Serafeim and Yoon (2021) also indicate a link between positive ESG news and an increase in stock prices, and conversely, negative ESG news is connected with a drop in stock prices. In addition, Hammami and Hendijani Zadeh (2019) also find that public exposure from the news is one of the main drivers of ESG transparency that can help mitigate the information asymmetry between the companies and their stakeholders.

Considering the considerable influence of news, this study seeks to investigate the analysis of news as a means to evaluate the ESG performance of companies. However, due to the fact that not all news articles pertain to the subject of ESG, a careful selection must be made. In this context, I delve into the potential of Generative Pre-trained Transformer (GPT) models to perform multi-class classification tasks. In 2020, OpenAI released the largest language model at that time, GPT-3, with 175 billion trainable parameters (T. B. Brown et al., 2020). The model has been trained from various sources, including CommonCrawl, with a total of 499 billion tokens (T. B. Brown et al., 2020). The paper shows that the GPT-3 model performed better on the LAMBADA dataset test than the state-of-the-art of that time, Turing-NLG, with an accuracy of 72.5% (Zero-Shot) to 86.4% (Few-Shot).

In late 2022, OpenAI introduced ChatGPT, marking a significant stride in the progression of NLP and Artificial Intelligence (AI). Originating from the GPT-3.5 series, ChatGPT represents a refinement and advancement beyond its predecessor, GPT-3 (OpenAI, 2022). Not only does it generate

responses that mimic human interaction, making it suitable for conversational applications, but it has also showcased remarkable capacities in technical fields. Notably, ChatGPT was reportedly capable of securing an entry-level software engineer position at Google and successfully navigating exams for law and business schools (Elias, 2023; Kelly, 2023). The substantial growth of ChatGPT is evidenced by its rapid user adoption, exceeding 100 million active users within two months post-launch and thus joining the ranks of the fastest-growing consumer applications in history, second only to Threads by Meta in the speed of user acquisition (Gordon, 2023).

Motivated by the surging interest in OpenAI's GPT models, I am drawn toward leveraging them to dissect and analyze the expansive and unstructured data within the news. This paper introduces a streamlined machine learning pipeline, employing a variety of models to ensure a smooth and intuitive measurement process. The initial phase involves collating the titles and leading paragraphs of news articles from the New York Times. Subsequently, named-entity recognition (NER) is implemented to isolate news that pertains to specific companies. Following this step, extract and categorize the news articles by using the text classifier. Lastly, sentiment analysis is applied to assign an ESG score to each company within each category. The main contributions of this thesis encompass:

1. A manually labeled, fine-grained ESG dataset comprising 4,500 news items.
2. An exploration into the capabilities of GPT models, evaluating their proficiency in performing classification tasks specifically within the ESG context.
3. A fine-tuned model based on the GPT model to enhance its functionalities in performing multi-class classification tasks in the ESG context.
4. A proposal of a systematic machine learning pipeline, which has been designed to assess companies' ESG performance utilizing information derived from news articles.
5. A thorough analysis of several companies' ESG performance, drawing from news articles to illuminate their activities and impacts in this domain.

The results of this study reveal that a fine-tuned model, which is less expensive and smaller in scale, can outperform the zero-shot prompting capabilities of a larger, more costly model in ESG classification tasks. Additionally, the machine learning pipeline used in this research demonstrates an ability to grasp the nuances and assess the sustainability performance of companies, as depicted in media reports. The analysis indicates that media attention varies across industries, resulting in a more thorough assessment of some companies over others. It is also worth mentioning that the environmental initiatives undertaken by companies are often underrepresented in media coverage, with a greater focus observed on governance and social aspects. Nevertheless, the findings from the machine learning pipeline also uncover areas where companies receive criticism from the media, pointing to opportunities for improvement in specific aspects.

The remainder of this paper is organized as follows. Section 2 provides an overview of the existing research on large language models and their applications in the ESG field. Section 3 outlines the research methods employed in this study. Section 4 portrays the findings of the study and the outcomes of the study, focusing on how companies are scored using the proposed approach. Section 5 goes into a detailed discussion of the findings and explains the limitations of the study. Finally, Section 6 concludes the paper with the conclusions.

## 2. Literature Review

### 2.1. Large Language Models

Over the past few years, there has been a significant shift in the field of NLP with the introduction of large language models (LLMs). These models are pre-trained foundational models that are self-supervised and are trained on extensive text datasets (Sejnowski, 2023; Sun et al., 2023). These LLMs, as described by A. H. Huang et al. (2023), use contextualized embeddings[1] that can represent a word with various vectors based on its context. This approach is a notable progression from the earlier word embedding models like word2vec and GloVe, where each word was tied to a single, fixed vector without taking into account the surrounding text (A. H. Huang et al., 2023). In contrast, LLMs adjust the vector representation of words based on their immediate textual context, making them especially effective in interpreting homographs[2] (A. H. Huang et al., 2023). As a result, LLMs offer more nuanced and accurate interpretations when analyzing texts compared to their predecessors.

Although the history of language models can be dated back more than 100 years, with early concepts being explored by Markov and Shannon based on probability theory (H. Li, 2022), a game-changing development came with the introduction of the transformer model by Vaswani et al. (2017). Originally designed for translation tasks, the transformer's defining idea is its self-attention mechanism, which allows the model to weigh the importance of various parts of an input sequence when producing an output (Douglas, 2023). Owing to its exceptional capabilities in language representation, transformer architecture has become foundational in almost all of today's pre-trained language models (H. Li, 2022).

The transformer architecture quickly gained popularity among researchers and split into three main categories: bidirectional, unidirectional, and sequence-to-sequence (Douglas, 2023; H. Li, 2022). First, there is a bidirectional model

[1] Embeddings are mathematical representations of objects or values, such as text, images, and audio, transformed into vectors based on their characteristics, attributes, and categorical associations.

[2] A homograph is a word that is spelled identically to another word yet differs in meaning.

using the encoder-only architecture like BERT (short for Bidirectional Encoder Representations from Transformers) (Devlin et al., 2018). BERT learns by hiding random words in a sentence and guessing them using the surrounding words for clues, which helps it get better results (Devlin et al., 2018; Douglas, 2023). Then, there is the unidirectional type, like GPT, that uses the decoder-only architecture (Radford et al., 2018). This architecture allows GPT to learn effectively from unlabeled data, which is helpful when there is no labeled data available (Radford et al., 2018). Finally, there is the sequence-to-sequence model, which uses both an encoder and a decoder. Some examples of this are BART, which stands for Bidirectional and Auto-Regressive Transformers (Lewis et al., 2019) and T5 (Raffel et al., 2019).

There are a lot of applications of LLMs across various industries and use cases. In healthcare, LLMs help answer medical questions. For example, Venigalla et al. (2022) create BioMedLM, a GPT model trained on biomedical data. Singhal et al. (2022, 2023) develop Med-PaLM and Med-PaLM 2, based on a pathways language model (PaLM), and Y. Li et al. (2023) make ChatDoctor, refining the large language model meta-AI (LLaMA) with a massive collection of dialogues between patients and doctors. In finance, Wu et al. (2023) introduce BloombergGPT, an LLM that is trained on a wide spectrum of financial data. In the legal world, Chalkidis et al. (2019) use LLMs to predict case outcomes from the European Court of Human Rights (ECHR), and Peric et al. (2020) modified a GPT-2 model to help write legal opinions.

Another fascinating area of research about LLMs is prompt engineering, which focuses on how to communicate with these models to get the best results effectively. While there are numerous plausible techniques for prompting, the field of prompt engineering still necessitates extensive experimentation to understand and perfect these methods (Kaddour et al., 2023). There is limited theoretical knowledge about why certain ways of wording a task work better, except for the fact that they produce superior practical outcomes (Kaddour et al., 2023). Various approaches have been investigated in prompt engineering, such as in-context learning, which enables LLMs to pick up new tasks without altering any parameters. There is also multi-turn prompting, which links a series of prompts and responses in a sequential manner, and chain-of-thought (CoT) prompting, which builds few-shot prompts through a sequence of intermediate reasoning steps leading to the final result, among other methods.

This study explores a prompt engineering technique by Sun et al. (2023). They introduce the 'Clue And Reasoning Prompting' (CARP) method, which adopts a step-by-step reasoning process designed for the intricate language patterns seen in text classification. It starts by nudging the language models to identify key clues. Next, it requests detailed reasoning, which ultimately helps in making the final classification. This technique impressively set a new state-of-the-art (SOTA) performance on four out of the five most common text classification benchmarks. Prompt engineering proves to be a valuable strategy when faced with constraints like limited datasets and computational resources for training or fine-tuning, managing to deliver results that can compete with more resource-intensive methods.

### 2.2. Large Language Models for ESG

Machine learning, in general, has been a key tool in ESG analysis and research. It has been used in a variety of ways, from shaping trading strategies to assessing risks. For instance, Chen and Liu (2020) use machine learning to understand a company's ESG premium and capture the ESG alpha to build an automatic trading strategy. Impressively, their approach is proven to outperform the NASDAQ-100 and S&P 500 indexes over a ten-year test period. On another front, Nguyen et al. (2020) tap into machine learning to predict a company's future emissions, which can help policymakers decide where to focus their efforts. Guo et al. (2020) look at ESG news to predict the volatility of stock prices.

Krappel et al. (2021) provide a heterogeneous ensemble to predict the ESG ratings of companies even if the company does not disclose its sustainability report. Furthermore, Polignano et al. (2022) and H. Kang and Kim (2022) utilize NLP to extract information from corporate sustainability reports and propose a machine learning pipeline to analyze the sustainability performance of companies. Incorporating news into the model, Borms et al. (2021) share a method to summarize ESG-related news using word patterns. They create a comparison using news-based ESG signals with the scores from an external data provider. Goutte et al. (2023) delve into how news sentiment, especially when it concerns ESG topics, can predict stock returns using data provided by the Global Database of Events, Languages, and Tone (GDELT) project.

The development of LLMs then further improves the nuance and comprehensiveness of textual data from different sources, such as reports or news articles. Different applications and improvements from the foundational models (e.g., BERT & GPT) have been explored by multiple researchers. Using a question-and-answering approach, Luccioni et al. (2020) develop ClimateQA using the BERT model, which assists in the analysis of financial reports regarding climate-relevant sections. Webersinke et al. (2021) build a language model, ClimateBERT, that was pre-trained on a dataset of over two million climate-related paragraphs, which improved the predictive performance in three climate-related tasks: text classification, risk and opportunity analysis, and fact-checking climate-related claims. This model was further applied to assess the disclosure of climate-related financial risks by Bingler et al. (2022).

Many researchers delve into topic modeling by leveraging LLM to extract ESG information. Mehra et al. (2022), for example, successfully fine-tune BERT using ESG texts from guides, case studies, blogs, reports, and other knowledge bases from the Accounting for Sustainability project. The fine-tuned model, called ESGBERT, is then applied to do classification tasks on the 10-Q filings of companies. Similarly, J. Lee and Kim (2023) make use of a BERT-based model to create a classification model based on Korean companies' sus-

tainability reports. Their model can sort sentences into four ESG categories, including those that are not relevant. They also tested their model on other types of data, such as news articles, to demonstrate its adaptability.

Aue et al. (2022) introduce a technique to predict ESG ratings by looking at patterns over time based on news articles. To identify news related to ESG and figure out the sentiment, they utilize BERT in their method. Fischbach et al. (2022) develop a model based on BERT to sort news headlines into three ESG categories and label irrelevant news, called ESG-Miner. While the model is good at identifying headlines related to the environment, its accuracy in determining ESG relevance could use some refinement. Nevertheless, specialized ESG classifier models are still developing, especially for fine-grained ESG topics. Tackling fine-grained ESG topics presents a significant challenge largely due to the absence of a universally accepted standard to define classes under the three ESG pillars. ESG rating institutions, such as MSCI, Refinitiv, and Sustainalytics, employ different methodologies to define the fine-grained classes.

The opportunity to further explore text classification within this domain remains substantial, considering there are only a handful of available models, and they present differing categorizations for the fine-grained classification. For instance, Nugent et al. (2020) pre-train a BERT model on financial news articles from the Reuters News Archive. They then used this trained model to sort texts into 20 different ESG topics and also to label them based on the United Nations Sustainable Development Goals (UN SDGs). H. Lee et al. (2023) use a different set of classes consisting of 35 ESG key issues by MSCI ESG rating guidelines. In their study, they utilize various models based on the BERT model to do the classification tasks.

Interestingly, the categories used by Nugent et al. (2020) and H. Lee et al. (2023) are divergent from those introduced by A. H. Huang et al. (2023). In a later study, A. H. Huang et al. (2023) create FinBERT, which is a version of BERT designed specifically for financial topics. They further provide several models for different purposes, such as FinBERT-tone for sentiment analysis of financial texts, FinBERT-ESG for ESG topic sorting, and FinBERT-ESG-9-categories for fine-grained ESG classification. In their detailed classification model, A. H. Huang et al. (2023) use 14,000 manually annotated sentences from ESG reports and annual reports to fine-tune FinBERT. This model can sort texts into nine distinct ESG categories: Climate Change, Natural Capital, Pollution and Waste, Human Capital, Product Liability, Community Relations, Corporate Governance, Business Ethics and Values, and Non-ESG.

Throughout this chapter, it is evident that a significant portion of research tends to favor the use of the BERT model for various applications, especially for fine-grained ESG classification. Despite this predominant focus on BERT, there exists a promising avenue of exploration in understanding and leveraging GPT models, particularly in the context of ESG topic identification from news articles. My aim is to contribute meaningful insights and perspectives on how GPT models can also be a good choice for this kind of task and provide a balanced view of the capabilities of different language models in handling the extraction of ESG-related information from textual data.

## 3. Methodology

In this section, I describe the proposed streamlined machine learning pipeline for measuring the sustainability performance of a company. It starts with the data collection from the news outlet to scoring the performance based on news sentiment analysis. I am using available machine learning models for NER tasks and sentiment analysis tasks, while I am using a fine-tuned GPT model for the fine-grained ESG classification tasks. The pipeline is shown in Figure 1. The pipeline is available in the Python language because there are models and libraries available to use for NLP tasks in this language.

### 3.1. Data Collection

This study extracts news articles from the New York Times, utilizing the Archive API. Recognized globally and based in New York City, the New York Times is a daily newspaper renowned for its wide-reaching influence and significant subscriber base, with over nine million subscribers worldwide (The New York Times Company, 2023). Besides its reputation and credibility to ensure the reliability and validity of the news data, the New York Times is chosen because it provides an API that allows for systematic data retrieval, unlike many other news outlets, which is pivotal for consistent and reproducible research practices. Furthermore, the availability of this API service free of charge aligns well with budgetary constraints and enables extensive data extraction without incurring additional costs. A collection of 1.8 million news articles, spanning from 2003 to 2022, forms the basis of data to be analyzed further. The chosen date range ensures that a substantial amount of data is available for analysis, providing a robust dataset to identify and analyze trends or patterns related to ESG performance over an extended period.

In alignment with the study's objectives, specific elements of the metadata—specifically headlines, lead paragraphs, web URLs, keywords, and published dates—were extracted for further analysis. The analytical focus of this study targets the combined text of headlines and lead paragraphs. Integrating both headline and lead paragraph ensures a coherent message is conveyed, as a headline or a lead paragraph alone often lacks the necessary context to fully understand the unfolding events. Keywords serve as a preliminary filter to isolate news pertaining to specific companies, streamlining the process of data analysis. Meanwhile, the publication dates provide a temporal framework to understand the sequence and timing of events.

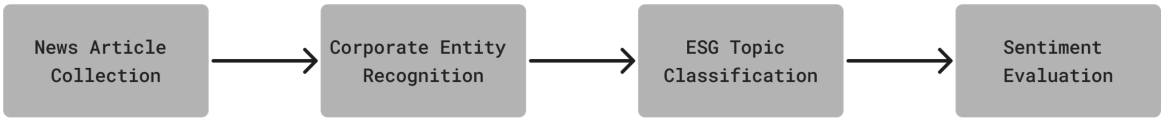


**Figure 1:** Machine Learning Pipeline for ESG Assessment

### 3.2. Company Selection and Named-Entity Recognition

In order to see the pipeline in action, this research selects eight companies to be examined. There are three criteria that are put in place to select the companies: market capitalization, sector representation, and news volume. All chosen companies are recognized as some of the largest entities within the S&P 500 based on their market capitalization to ensure it encompasses firms with significant impact and influence in the market. The companies were also selected to provide representation from diverse market sectors, giving breadth to the study by allowing for insights to be derived across different industries. Finally, companies with a more substantial presence in the news were prioritized, as a higher volume of available data enables a more in-depth and viable analysis.

The eight selected companies come from three different sectors: the technology/communication services sector (technology), consumer staples, and healthcare. Apple, Microsoft, Google, and Meta were chosen from the technology sector due to their gigantic market capitalization. Apple, Microsoft, and Google are the biggest companies within the S&P 500, with a total of over US$ 7 trillion in market capitalization (Johnston, 2022). Although Meta is not in the top four, it is included due to its significant presence in news articles and its operational similarities with the other three technology giants. It was challenging to select companies in the consumer staples and healthcare sectors because the amount of news for these sectors is not as many as in the technology sector. Therefore, I decided to select Coca-Cola and Pepsi for the consumer staples industry and Pfizer and Johnson & Johnson (J&J) for the healthcare industry because they tend to be more frequently featured in the news compared to other companies in their respective field.

It is worth mentioning that the list of keywords from the New York Times can sometimes include items that may not be directly related to the news article in question. Take the keyword "Apple Inc." as an example, and there could be instances where this is included even if the news piece primarily discusses only Microsoft's activities. To avoid this kind of impreciseness and to ensure the relevance of the company to the article, it is important to deploy a mechanism for accurate company name detection. In addressing this challenge, I utilize NER, which is notable for its capabilities to identify and categorize entities within the text into specified categories, including organizations, persons, and locations, among others. By applying NER, it is possible to extract the company of interest from the text, ensuring that the company is relevant to the article.

This research applies NER via the spaCy[3] library, specifically using the 'en_core_web_md' package, version 3.6.0, which is the medium-sized English model proficiently trained on web text, including news articles. This model is chosen for its capability to swiftly perform statistical entity recognition, such as identifying various types of named and numeric entities within text. The names of firms can be selectively extracted by targeting the label recognized as 'ORG'. Following the extraction, a string-matching method is applied to further refine the collection process based on the company of interest.

### 3.3. Text Classification

This study emphasizes the exploration of the capabilities of GPT models, particularly in executing ESG classification tasks. Therefore, unlike other machine learning tasks, including NER and sentiment analysis, this machine learning pipeline uses the results from this exploration. In this section, I outline the methods employed to investigate their potential to achieve fine-grained ESG classifications, such as the definitions of each class used in the model, the dataset for training and validating the model, and the procedure for fine-tuning the GPT models and validating zero-shot classification.

#### 3.3.1. Class Definitions

The ESG topics used in this paper are inspired by MSCI ESG Key Issues (MSCI, 2023b) with minor adjustments. MSCI ESG Key Issues were selected as the basis for the class definitions in this study mostly due to its credibility, which makes it used by the majority of ESG exchange-traded funds (ETFs) (Hirai et al., 2021). The eight ESG topics are *Climate Change, Resource Stewardship, Environmental Opportunities, Human Capital, Product Liability, Social Opportunities, Corporate Governance,* and *Business Ethics.* One other class is *Non-ESG* to flags irrelevant news articles. The definition and example of each class are described as follows.

*Climate Change*

Pertains to news articles discussing topics such as carbon emissions reduction initiatives, the carbon footprint of products, climate change vulnerabilities, and financial initiatives or instruments designed to mitigate the impact of climate change. This can include policies, new technologies, or corporate strategies targeting climate change. The examples are as follows.

[3] From Honnibal and Montani (2017)

- *CEO of UK-based energy supplier Drax shares how the company, formerly 100% reliant on coal, reduced its carbon emissions by 85%. The company now has ambitions to not just be carbon neutral, but carbon negative.*
- *ExxonMobil Formally Joins The Net-Zero By 2050 Bandwagon. U.S. energy giant ExxonMobil announced Tuesday a formalized plan to cut its scope 1 and scope 2 carbon emissions to "net-zero" by the year 2050.*

*Resource Stewardship*

Involves articles highlighting how companies manage natural resources and waste. It encompasses a range of issues, including but not limited to water conservation, biodiversity, sustainable land use, responsible raw material sourcing, toxic emissions reduction, and effective waste management, including electronic waste. The examples are as follows.

- *Starbucks to Offer Reusable Cups in All EMEA Stores by 2025. Starbucks Corp will offer reusable cups in stores across Europe, the Middle East and Africa by 2025 in an effort to reduce the amount of single-use waste heading to landfill.*
- *The Coca-Cola Company And The Ocean Cleanup Join Forces To Clean Up 15 Of The World's Most Polluting Rivers Of Plastic Waste. The Coca Cola Company and The Ocean Clean-Up project have announced they will be collaborating on a ground-breaking partnership to clean up some of the world's worst polluting rivers - and collect plastic waste which can be recycled to make new bottles.*

*Environmental Opportunities*

Includes articles that focus on the potential opportunities arising from environmental conservation efforts. It covers green technology innovations, renewable energy initiatives, sustainable building projects, and financial investments targeting environmental sustainability. The examples are as follows.

- *China Clean Energy Giants Unveil World's Largest Wind Turbines. Ming Yang Smart Energy Group Ltd. unveiled the world's largest wind turbine, an offshore behemoth whose more than 140-meter-long blades will sweep across an area larger than nine soccer pitches.*
- *Honda Recommits To Fuel Cells As It Looks For New Markets. Honda is planning to offer up its new generation fuel cell systems for commercial vehicles, construction equipment and stationary power systems beginning in 2025.*

*Human Capital*

This class encompasses news items related to labor management, employee welfare, and workforce development. It can include articles about health & safety protocols, human capital development programs, and supply chain labor standards, including diversity, equity, and inclusion initiatives in the workplace. The examples are as follows.

- *Female employees file class-action discrimination suit against Black News Channel. Thirteen women who worked at Black News Channel say in lawsuit they were paid less than men and disciplined for being too aggressive in the workplace.*
- *Amazon warehouse workers suffer muscle and joint injuries at a rate 4 times higher than industry peers. Amazon workers are four times as likely to incur strains, sprains and other repetitive stress injuries as workers in non-Amazon warehouses.*

*Product Liability*

Relates to articles discussing aspects of product safety and quality, including chemical safety and consumer financial protection. It includes topics such as privacy and data security issues and socially responsible investment, which might impact product liability. The examples are as follows.

- *Google is facing a lawsuit after a privacy flaw in its contact tracing tech exposed Android users' data to third-party apps. The lawsuit alleges that Google exposed participants' private personal and medical information when they opted into using contact tracing apps.*
- *Nestle recalls more than 760,000 pounds of Hot Pockets because they might contain bits of plastic and glass. Four people contacted Nestlé when they found "extraneous material" in their Hot Pockets, the USDA said.*

*Social Opportunities*

This category covers articles focusing on the societal benefits generated through corporate initiatives. It includes news on community financing, enhancing healthcare access, nutrition and health opportunities, educational initiatives, and investments aimed at social development. The examples are as follows.

- *HBCUs Team Up With Wells Fargo To Improve Financial Wellness For College Students Of Color. Wells Fargo is investing $5.6 Million in a financial literacy and wellness program designed for college students of colorant their surrounding communities.*
- *DoorDash Establishes Grant Program For Women- And BIPOC-Owned Restaurants Disproportionately Affected*

*By Covid. DoorDash's Main Street Strong Accelerator provides financial support and specialized educational resources specifically to women-, immigrant- and people of color-owned businesses that have been disproportionally impacted by the Covid-19 pandemic.*

*Corporate Governance*

Deals with articles related to the structural and strategic management aspects of corporations. This includes topics such as ownership and control dynamics, board composition and performance, executive remuneration, accounting transparency, and significant executive team changes, including the hiring or resignation of C-level executives and directors. The examples are as follows.

- *Black employees are questioning Peloton about their pay, as the fitness giant's CEO pulls in a $17.8 million compensation package. Black employees are questioning Peloton execs about pay, while the CEO earned $17.8 million last fiscal year and the median employee earned $56,084.*
- *Discord Adds Ex-Netflix, Block Executives To Board Ahead Of Possible IPO. The two C-suiters' experience at public companies is the latest indication that Discord intends to go public soon.*

*Business Ethics*

Encompasses articles on the ethical considerations of business operations. It includes subjects such as tax transparency, anti-corruption measures, fraud prevention, and adherence to ethical business practices and regulations. The examples are as follows.

- *Former Netflix executive convicted of fraud after orchestrating more than $500,000 in bribes and kickbacks. As Netflix's IT chief, Michael Kail approved contracts with tech startups in exchange for kickbacks, even buying a house with the funds, a jury found.*
- *McDonald's to pay France $1.3 billion in tax fraud case. McDonald's France and related companies have agreed to pay $1.3 billion to the French state to settle a case in which the fast-food giant was accused of vast tax evasion*

*Non-ESG*

This class is for articles that do not fit into any of the above ESG categories. It can include a wide range of topics not directly related to environmental, social, and governance issues. The examples are as follows.

- *McDonald's temporarily removes the Chicken Big Mac from menus, saying that it's struggling to keep up with demand. The limited-edition burger was surprisingly popular, which led to its removal from UK restaurant menus while stocks are replenished.*
- *Ferrari Sparkling Wine Becomes The Official Toast Of Formula One. Italian sparkling wine will be at the podium.*

3.3.2. Dataset

For training and validating purposes, I used a different source of news articles than what has been described in Chapter 3.1 to minimize bias during the later application. In this step, I used the dataset from the GDELT project as the source of my news data. This project monitors news media from all over the world in over 100 languages (The GDELT Project, n.d.-a). It is considered one of the largest and most comprehensive open databases created by Kalev Leetaru in a research collaboration with a lot of institutions, including, among others, Google, JSTOR, and the Internet Archive (The GDELT Project, n.d.-b). It now consists of over a quarter-billion event records in over 300 categories covering events from 1979 to the present (The GDELT Project, n.d.-b).

GDELT offers different datasets for different purposes, such as GDELT Event Database, GDELT Global Knowledge Graph, and GDELT Article List. The most comprehensive database is the GDELT Event Database, which contains historical data from 1979 to date. However, for the purpose of this thesis, this database does not provide information on similar data as the lead paragraph from the New York Times API or a summary of the news articles. Only the GDELT Article List database provides this kind of information. Unfortunately, this database does not incorporate historical news, as it only started to collect the data in January 2020. Another challenge to using this database is the missing data from the news summary, which was only available in late 2020. This poses an obstacle where the usable data is only in a limited timeframe.

I have gathered news from several prominent media outlets to address the timeframe limitation and guarantee a sufficient volume of articles. These include the Wall Street Journal, Los Angeles Times, Bloomberg, Business Insider, the Economist, Forbes, the Washington Post, and ESG News. The collection spans from 2021 to April 2023, containing a total of 698,433 news pieces. These articles will then be selected and manually annotated to each ESG topic defined in the previous section. Nevertheless, the dataset has a high level of duplicate news articles, which needs to be removed. Moreover, there are a lot of irrelevant news articles because the dataset contains articles for all topics, including sports, arts, and others.

The duplicated news articles and unrelated ones are removed. Precisely, I aimed to include articles from the business, science, technology, and environment if it is identifiable through their URL structures. Given the vast amount of data, manually categorizing each article is daunting. To streamline this process, I initially employed the FinBERT ESG 9 Categories model by A. H. Huang et al. (2023). This model

was particularly useful for basic categories like environmental, social, governance, and non-ESG. However, it was not foolproof. Some articles were not correctly sorted, and given our unique classification requirements, I still had to manually review and annotate many pieces to ensure their accurate categorization. This blend of automated and hands-on approaches allowed me to generate a more reliable dataset while saving time.

At first, I planned to add the stakeholder opposition topic and split the resource stewardship topic into natural capital and pollution & waste. However, finding news articles related to these topics in the dataset was challenging. Therefore, I excluded the stakeholder opposition and combined both the natural capital and pollution & waste topics into a broader resource stewardship category. Ultimately, I generated a manually annotated dataset with 4,500 articles in total, ensuring 500 articles for each topic. I then divided this dataset into training and validation sets using an 80:20 split.

#### 3.3.3. Methods

This study investigates GPT models' capabilities in handling multi-class classification tasks, focusing on two distinct approaches: fine-tuning and zero-shot prompting. While fine-tuning refines a previously trained model for a particular task, zero-shot prompting relies on the model's inherent knowledge, eliminating the need for further training. GPT models, as large language models, possess a vast amount of knowledge, allowing them to make predictions, even for unfamiliar tasks. The fine-tuning approach is advantageous for models with a narrower knowledge base and is also cost-effective. Conversely, zero-shot prompting proves valuable when working with a sparse dataset.

This paper examines different models for fine-tuning and zero-shot prompting. For the fine-tuning method, I delve into the capabilities of the 'ada' model from the GPT-3 series. Being both cost-effective and fast, 'ada' emerges as a fitting choice for fine-tuning. Meanwhile, for zero-shot prompting experiments, I employ the 'gpt-3.5-turbo-0613' from the GPT-3.5 lineup. Its vast knowledge base makes it well-suited for such experiments. These model selections aim to optimize the balance between performance and efficiency. I hope the comparison of these two distinct approaches sheds light on the potential and limitations of each method.

For the fine-tuning method, OpenAI has furnished users with tools to streamline the process. While the GPT-3 model remains a closed-source offering, direct fine-tuning on personal hardware is not feasible. Consequently, I had to rely on OpenAI's dedicated services to carry out this task. There is a prerequisite to prepare the dataset in a specific manner. Particularly, it should comprise the 'prompt,' which serves as the input, and the 'completion,' which denotes the expected output. This dataset is to be formatted in the JSONL structure.

I conducted a series of experiments to achieve the optimal configuration for the fine-tuned model. There are two key techniques: the inclusion of a prompt in the input and the application of dummy label[4] in the output. This led to four distinct combinations to test:

(a) A version without any prompt and using the original label

(b) A version without any prompt but introducing the dummy label

(c) Incorporating the prompt while sticking to the original label

(d) Incorporating the prompt but introducing the dummy label

The pairing of the original label with the dummy label I used are detailed in Table 1. To provide a clearer understanding of these configurations, I illustrate the examples in Table 2.

In the zero-shot prompting method, I followed the approach of Sun et al. (2023) called CARP. However, CARP's initial tests were focused on sentiment analysis. In this study, I aim to evaluate how CARP performs when applied to multi-class classification tasks, particularly for ESG topics. Due to the absence of a standard definition for fine-grained ESG topics and to make sure the CARP output aligned with the objective of this paper, I included the definition of each class in every prompt. This ensures the model uses the exact definitions I used during labeling and only produces classes from the list. I split the prompt into two sections: system and user. In the system content, I prompted the model to act as an ESG news articles classifier and gave it the class definitions. In the user content, I structured the prompt following the CARP method.

#### 3.3.4. Sentiment Analysis

The concluding step of this process involves analyzing the sentiment of the news articles and categorizing them as negative, neutral, or positive. There are many machine learning models designed for sentiment analysis tasks. To find the best fit, I chose four models tailored to financial news or texts, given their similarity to the dataset in this study. Precisely, I selected the top three models trained on the 'financial_phrasebank' dataset (Malo et al., 2013) based on accuracy rankings from the Papers With Code website (Papers With Code, n.d.), along with the sentiment analysis model from FinBERT by A. H. Huang et al. (2023). The three models trained on the 'financial_phrasebank' dataset are Sigma/financial-sentiment-analysis[5] (Sigma), Farshid/bert-large-uncased-financial-phrasebank-allagree2[6]

[4] Labeling data with random strings or placeholders, often referred to as 'metasyntactic variables,' is a technique used to mitigate bias in language models.

[5] https://huggingface.co/Sigma/financial-sentiment-analysis

[6] https://huggingface.co/Farshid/bert-large-uncased-financial-phrasebank-allagree2

**Table 1:** Pairing Between Original and Dummy Labels

| Original Label | Dummy Label |
|---|---|
| Climate Change | baz |
| Resource Stewardship | qux |
| Environmental Opportunities | roc |
| Human Capital | tuv |
| Product Liability | dap |
| Social Opportunities | stu |
| Corporate Governance | klo |
| Business Ethics | xya |
| Non-ESG | nop |

**Table 2:** Examples of Input and Output

| Model | Input (Prompt) | Output (Completion) |
|---|---|---|
| (a) | CEO of UK-based energy supplier Drax shares how the company, formerly 100% reliant on coal, reduced its carbon emissions by 85%. The company now has ambitions to not just be carbon neutral, but carbon negative. -> | Climate Change |
| (b) | CEO of UK-based energy supplier Drax shares how the company, formerly 100% reliant on coal, reduced its carbon emissions by 85%. The company now has ambitions to not just be carbon neutral, but carbon negative. -> | baz |
| (c) | Classify the following text into one of the following classes: [' Climate Change', ' Resource Stewardship', ' Environmental Opportunities', ' Human Capital', ' Product Liability', ' Social Opportunities', ' Corporate Governance', ' Business Ethics', ' Non-ESG'] Text:\n'"CEO of UK-based energy supplier Drax shares how the company, formerly 100% reliant on coal, reduced its carbon emissions by 85%. The company now has ambitions to not just be carbon neutral, but carbon negative.'" -> | Climate Change |
| (d) | Classify the following text into one of the following classes: [' baz', ' qux', ' roc', ' tuv', ' dap', ' stu', ' klo', ' xya', ' nop'] Text:\n'"CEO of UK-based energy supplier Drax shares how the company, formerly 100% reliant on coal, reduced its carbon emissions by 85%. The company now has ambitions to not just be carbon neutral, but carbon negative.'" -> | baz |

(Farshid), and mrm8488/distilRoberta-financial-sentiment[7] (mrm8488). All models are available on HuggingFace.

In determining the optimal model from the four contenders, I did a systematic and multi-step evaluation process. To begin, a preliminary assessment was conducted by evaluating the results from each model using the news articles from the New York Times to assess Apple's performance. I manually annotated 50 news articles to serve as ground truth for this evaluation. As a result, two of the four models showcased similar accuracy rates, which are FinBERT and the mrm8488's model, making them the frontrunners. Next, to differentiate the two top-performing models, a confusion matrix was examined, as shown in Table 3.

One notable observation from this matrix is that, in instances where sentiments are polar opposites, FinBERT frequently classified articles as negative, whereas the counterpart model leaned towards a positive classification. Given this intriguing divergence in sentiment classification, a focused annotation was undertaken. A subset of 25 articles, for which FinBERT predicted a negative sentiment and the alternative model predicted a positive, were manually annotated to understand the models' behavior better. After the detailed assessment, it was deduced that the FinBERT model provided more accurate and suitable classifications for the task in this study.

The sentiment conveyed in the news articles serves as the foundation to develop the company's performance score. For quantifying this sentiment score within each category, distinct weights are allocated to each sentiment type. Specifically, a weight of -1 is designated to negative sentiment, 0 to neutral, and 1 to positive sentiment. The cumulative score is derived by multiplying the total count of articles within each sentiment category by its respective weight. Given the potential for varied data distribution across categories and companies, this raw score is then normalized. This is achieved by dividing the cumulative score by the total sentiment count, ensuring a more equitable comparison of ratings across classes and companies. The formula for the normalized is described

[7] https://huggingface.co/mrm8488/distilroberta-finetuned-financial-news-sentiment-analysis

**Table 3:** Confusion Matrix of the FinBERT and mrm8488 Model

| | | FinBERT | | |
|---|---|---|---|---|
| | | Negative | Neutral | Positive |
| **mrm8488** | Negative | 157 | 93 | 9 |
| | Neutral | 127 | 845 | 71 |
| | Positive | 65 | 125 | 156 |

in the Equation 1.

$$\text{Normalized Score} = \frac{[(-1 \times \text{number of negative sentiments}) + (1 \times \text{number of positive sentiment})]}{\text{total number of sentiments}} \quad (1)$$

## 4. Empirical Results

### 4.1. Text Classification Results

In a series of experiments of fine-tuning the 'ada' model, one of the GPT-3 base models, the same settings were applied to the four scenarios. I adhered to the default configurations suggested by OpenAI for various parameters, including the number of epochs, batch size, and learning rate multiplier. In particular, the default number of epochs in training is four, the batch size is set to approximately 0.2% of the number of examples in the training set (with a maximum limit of 256), and the learning rate multiplier defaults to 0.05, 0.1, or 0.2 depending on final batch size (OpenAI, n.d.). As the number of examples was identical in every scenario, sticking to the default settings ensured uniformity across all experimental setups. This uniformity is crucial for a fair comparison between different scenarios, making sure that any variations in performance are due to the treatment itself.

Each fine-tuning session in our experiments lasted around 25 minutes on average, while the cost for fine-tuning the base model was $0.0004 for every 1,000 tokens. Notably, in the first setup (a), where I used the original labels without any prompts, it took about 25 minutes and 8 seconds, and the cost for training on 793,328 tokens was $0.32. In the second setup (b), where I applied dummy labels without prompts, the process was slightly quicker and cheaper. It took around 24 minutes and 47 seconds, costing $0.31 to train on 780,528 tokens. The third and fourth setups, (c) and (d), which involved the use of prompts, did not necessarily take more time but were more expensive. For setup (c), the fine-tuning lasted about 24 minutes and 51 seconds, incurring a cost of $0.67 to train on 1,687,176 tokens. In the final setup (d), the duration was almost identical at 24 minutes and 49 seconds, but it was slightly cheaper at $0.62 for training on 1,559,176 tokens. The summary of the duration, cost, and number of tokens trained is presented in Table 4.

In the experiment where I employed the zero-shot prompting method with the CARP approach, it was sufficient to assess the outcomes using the same validation set as I used for the fine-tuning. This approach focused solely on the validation phase, eliminating the need to track the duration since there was no training phase involved. In general, for the GPT-3.5 turbo model utilized in this particular experiment, the cost structure was as follows: $0.0015 per 1,000 input tokens and $0.002 per 1,000 output tokens. In the process of validating a set of 900 examples, the total incurred cost amounted to $1.21, covering the processing of 806,073 tokens. The comparison of the cost and number of tokens processed between fine-tuning and zero-shot prompting is summarized in Table 4.

During the fine-tuning process, we noticed distinct performance trends for each setup. In the first setup, the accuracy started at 0.81 at the 900th step and increased significantly to 0.85 after the 2700th step. By the end of the observed range, at 3600 steps, the accuracy only improved slightly to 0.86, suggesting that the model might be reaching its optimal performance. In contrast, the second setup showed a steady climb in accuracy from 0.78 to 0.86, indicating consistent learning without any major fluctuations. The third setup began with an accuracy of 0.82 at the 900th step, slightly outperforming the first and significantly surpassing the second at similar intervals. This improvement persisted until the 2700th step, reaching 0.875 accuracy, but then plateaued. The fourth setup began similarly to the second, with a 0.79 accuracy at the 900th step. Yet, it experienced a notable rise to 0.84 by the 1800th step and gradually increased to 0.86 and 0.87 at the 2700th and 3600th steps, respectively. The accuracy trends for these setups can be seen in Figure 2–5.

From the results above, we can observe that using the original label gives better results after the first round of training. This could be because the GPT-3 model already understands these labels, while I used dummy labels in the second and fourth cases. However, a better performance in the first round does not guarantee the best final results. For instance, the first scenario actually had the lowest accuracy among all, even though it started strong. In terms of optimal stopping point, it is worth mentioning that in the third model, additional training beyond the point where it plateaus might be unnecessary. In contrast, the other three setups could still improve from further training, as they continue to show better results with each round of training.

To gain a well-rounded understanding of the model's performance, I leveraged training loss, accuracy, and weighted F1-score[8] to draw insights about the performance. In the first setup (a), the training loss was 0.021, and it achieved an accuracy of 0.862 with a weighted F1-score of 0.0862. The

[8] The weighted F1-score is a metric that assesses model accuracy by calculating and then averaging each class's F1-score (the harmonic mean of precision and recall) according to its prevalence in the dataset.

**Table 4:** Comparison of Cost and Number of Tokens Processed

| Scenario | Duration | Cost ($) | Tokens |
|---|---|---|---|
| (a) | 25 minutes and 8 seconds | 0.32 | 793,328 |
| (b) | 24 minutes and 47 seconds | 0.31 | 780,528 |
| (c) | 24 minutes and 51 seconds | 0.67 | 1,687,176 |
| (d) | 24 minutes and 49 seconds | 0.62 | 1,559,176 |
| CARP | n/a | 1.21 | 806,073 |

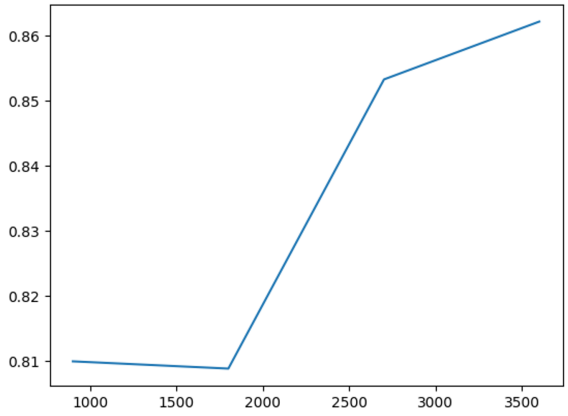


**Figure 2:** Accuracy and Steps Model (a)

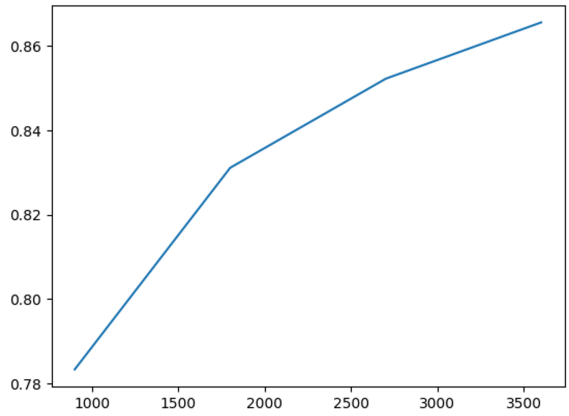


**Figure 3:** Accuracy and Steps Model (b)

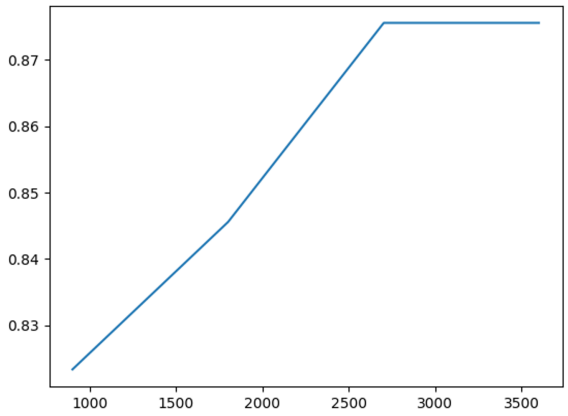


**Figure 4:** Accuracy and Steps Model (c)

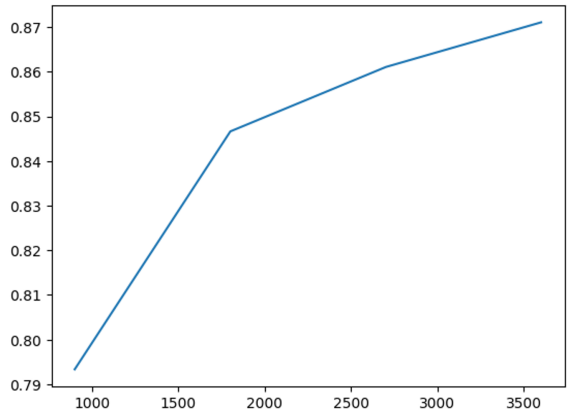


**Figure 5:** Accuracy and Steps Model (d)

second setup, setup (b), had a slightly higher training loss at 0.025 but managed to reach an accuracy and weighted F1-score of 0.865. Notably, the third scenario (c) demonstrated a superior efficacy with a lower training loss of 0.013 and an accompanying accuracy and weighted F1-score both at 0.875. Lastly, setup (d) had a training loss of 0.012, which is the lowest among the four, with slightly lower accuracy and weighted F1-score compared to (c), both at 0.871. From these findings, it becomes evident that setups (c) and (d) both exhibit low training losses, with setup (d) having the lowest. Low training loss is an indicator that the model has learned the underlying patterns from the data it was trained on. Furthermore, both (c) and (d) have high accuracy and weighted F1-score values, with (c) outperforming all. This suggests that model (c) is making correct predictions a high percentage of the time and is robust in terms of both false positives and false negatives.

From the zero-shot prompting experiment using CARP, it was observed that, from the 900-validation data, CARP achieved a moderate range of accuracy at 0.71 while the weighted F1-score was at 0.7. The similarity between these two values also indicates that it got balanced precision and recall. Compared to the results from fine-tuning, it is apparent that fine-tuning provides better performance. More consistency in both accuracy and weighted F1-score also indicates that the fine-tuned models are more discriminative, leading to fewer false positives and false negatives by the fine-tuned models. The superior performance of the fine-tuned models underscores the effectiveness of tailoring a model to a specific task, even though the pre-trained model

has fewer parameters. Despite potentially lower performance than fine-tuned models, the results from the zero-shot prompting experiment showcase the utility of this approach. Being able to achieve moderate accuracy without any task-specific training is still commendable and can be particularly useful in scarce labeled data scenarios, or rapid deployment of a model is necessary. Table 5 summarizes the performance results of the fine-tuned models and CARP experiment.

Upon analyzing the experimental results from fine-tuning and zero-shot prompting methods, a number of valuable observations come to light. For the fine-grained ESG topic classification tasks, a fine-tuned smaller model tends to outperform a larger model employing zero-shot prompting. Furthermore, the financial implications of training the smaller model are considerably less than those associated with validating the efficacy of zero-shot prompting. While OpenAI's pricing is marginally higher for fine-tuned models at $0.0016 per 1,000 tokens, the token consumption is reduced due to the absence of the need for repeated class descriptions in every prompt. Therefore, in an economical and efficient manner, fine-tuned models present a more viable option for this task compared to the large GPT-3.5 model. Evaluating the fine-tuned models, both (c) and (d) demonstrated impressive performance, marked by low training losses. To achieve optimal results, I prioritized accuracy and F1-score. As a result, model (c) was selected for the classification step of the machine learning pipeline proposed in this research.

### 4.2. Corporate Sustainability Performance Analysis

As detailed in Chapter 3, the process for analyzing news articles begins with gathering content from the New York Times through its Archive API. I then narrow down these articles, selecting only those relevant to the eight chosen companies, utilizing both keywords provided by the New York Times API and NER for effective filtering. Subsequently, I employ the FinBERT-ESG-9-categories model[9] to further sift through the data, setting aside irrelevant articles. This step is essential as it helps in omitting unnecessary content from the analysis and saves on potential costs that might have been incurred if the fine-tuned model were applied to irrelevant articles. Instead, the fine-tuned model is only applied to categorize news pieces into the other classes specified in this research. In the final stage, I conduct sentiment analysis on each article, aiming to gauge a company's performance based on the news content.

The evaluation process encompasses eight diverse companies, including technology giants Apple, Microsoft, Google (under Alphabet), and Meta (formerly Facebook), alongside major players in the beverages industry such as The Coca-Cola Company and PepsiCo, as well as pharmaceutical firms Pfizer and Johnson & Johnson. This varied selection allows for a comprehensive analysis across different sectors. I aim to gain insights into how the assessments vary from one industry to another, considering the unique characteristics of each sector. The number of news articles related to each of these companies is laid out in Table 6. The detailed assessment for each company is described in the following sections.

#### 4.2.1. Apple Inc.

Apple Inc. (Apple) is the largest company in the world in terms of market capitalization, having a staggering value of US$2.65 trillion as of 2022 (Johnston, 2022). Established in 1976 by Steve Jobs, Steve Wozniak, and Ronald Wayne in Los Altos, California, this American manufacturer has evolved to become a major player in the technology industry (Linzmayer, 2004). The company specializes in designing, manufacturing, and marketing a wide array of electronic devices and services, including smartphones, personal computers, tablets, wearables, peripherals, and various support services (Apple Inc., 2022). Some of its most famous products encompass the iPhone, Mac, and iPad. Apple's headquarters are located in California, and it employs approximately 164,000 full-time equivalent staff members (Apple Inc., 2022). Under the leadership of Chief Executive Officer (CEO) Tim Cook, who continued the reins from Steve Jobs in 2011 (Apple Inc., n.d.), Apple has continued to flourish and maintain its position at the forefront of technological advancement.

In our examination of 724 New York Times articles focused on Apple, the machine learning pipeline introduced in this study provided some insightful findings. Notably, Apple's societal impacts remain a central focus in media coverage. A dominant portion of the articles, amounting to 268, were classified under 'Product Liability,' reflecting a keen interest in the company's product offerings and related concerns. Additionally, governance, especially in terms of 'Business Ethics,' received significant attention with 237 articles. The 'Corporate Governance' category followed closely, comprising 107 articles, indicating a steady interest in Apple's corporate policies and practices. Environmental themes, on the other hand, were scarcely represented, with categories like 'Climate Change,' 'Resource Stewardship,' and 'Environmental Opportunities' garnering a combined total of just 12 articles. This limited environmental coverage suggests that the media might be underrepresenting Apple's efforts in this domain or that Apple's environmental initiatives were less newsworthy during the period of analysis.

Digging into the sentiment analysis of these articles reveals further insights into the media's perception of Apple's actions in each category. Overall, the sentiment expressed across the articles predominantly skews towards neutrality, with nearly 60% (428 out of 724) of the articles classified as such. This suggests a measured approach by the news outlet when reporting on Apple's multifaceted operations. In the 'Product Liability' theme, which held the largest share of articles, for example, a majority (153) maintained a neutral tone, while 86 leaned negative, and only 29 expressed a positive outlook. Moreover, in terms of 'Business Ethics,' the articles featured a neutral sentiment to a great degree (131), although negative perspectives were also evident in 81 news pieces. The remaining in this class were classified as positive. The 'Corporate Governance' subset showcased a largely neu-

[9] See A. Huang (2022)

**Table 5:** Comparative Evaluation Metrics of Different Models

| Scenario | Training Loss | Accuracy | Weighted F1-Score |
|---|---|---|---|
| (a) | 0.021 | 0.862 | 0.862 |
| (b) | 0.024 | 0.865 | 0.865 |
| (c) | 0.013 | 0.875 | 0.875 |
| (d) | 0.012 | 0.871 | 0.871 |
| CARP | n/a | 0.706 | 0.703 |

**Table 6:** Number of News Articles per Company Before and After Excluding Non-ESG Content

| Company | Number of News Articles | Without Non-ESG |
|---|---|---|
| Apple | 1,648 | 724 |
| Microsoft | 1,376 | 733 |
| Alphabet (Google) | 2,756 | 1506 |
| Meta | 2,113 | 1161 |
| Coca-Cola | 167 | 101 |
| Pepsi | 106 | 61 |
| Pfizer | 419 | 364 |
| Johnson & Johnson | 254 | 229 |

tral sentiment as well, accounting for 75 of its 107 articles. While environmental topics received limited attention, it is worth mentioning that articles focused on this aspect largely maintained a neutral to positive sentiment. The sentiment results for each category is described in Figure 6.

In terms of the scoring for Apple's performance using the sentiment, only the 'Environmental Opportunities' and the 'Social Opportunities' received a positive sentiment in general. Each of them got a total (normalized) score of 1 (0.2) and 4 (0.14), respectively. The 'Product Liability' and 'Business Ethics' aspects emerged as prominent areas of concern, with a total (normalized) score of -57 (-0.21) and -56 (-0.24). This indicates that despite a substantial volume of neutral coverage, the negative sentiment still slightly outweighs the positive. The critical stance is also echoed in the 'Human Capital' domain, with a total (normalized) score of -14 (-0.20). In regards to its performance in the 'Corporate Governance' realm, Apple hovers close to neutral with a modestly negative total (normalized) score of -2 (-0.02). Table 7 shows the comparison of scoring between the technology companies. The detailed scoring is available in the Appendix.

Diving deeper into the news articles, in regards to the 'Product Liability' aspect, there are various themes considered as negative. Concerns highlighted by media include privacy issues related to Apple's products and services, as well as the company's restrictive practices in managing its app store (e.g., Holpuch, 2022; Nicas, 2019). In terms of 'Business Ethics,' the discussion is often framed around antitrust concerns, with Apple's policies and fees for app developers being perceived as unfair (e.g., Nicas et al., 2020; Satariano, 2021). In the other facets, Apple was also scrutinized for its labor practices, where reports of stringent working conditions, particularly for workers in China, and friction with the labor unions (e.g., Gough and Chen, 2014; Zhuang, 2022). However, Apple has also gained positive attention for its initiatives aimed at enhancing healthcare, showcasing the company's concern for societal well-being (e.g., Singer, 2018). Despite this positive attention, it is noted that Apple has received less media attention compared to the more contentious topics.

According to MSCI's evaluation, Apple's approach to ESG matter is rated as moderate, placed in the middle range with a 'BBB' rating among 137 companies in the technology hardware, storage, and peripherals industry (MSCI, 2023a). When dissecting the components of the assessment, it is observed that Apple lags, particularly in areas concerning business ethics and labor practices within its supply chain (MSCI, 2023a). This viewpoint is echoed in the findings in the previous paragraph, where Apple is under significant scrutiny for ethical concerns and labor standards at its supplier locations. Interestingly, regarding issues of privacy and data security (part of product liability), MSCI categorizes Apple as average (MSCI, 2023a), contrasting with the more critical perspective of numerous negative sentiments identified in this research. At the same level, Apple is also considered average for its efforts in corporate governance and handling electronic waste (MSCI, 2023a). The present study parallels these findings, with the sentiment towards Apple's corporate governance being fairly mixed, albeit with a slight tilt toward the negative. Unfortunately, there were only a few articles talking about resource stewardship and none of them discussed Apple's waste management.

MSCI believes that Apple is a frontrunner compared to its peers in terms of human capital development and advancing clean technology. Despite Apple's efforts in these areas, they are still underreported by the news media. There are only a handful of articles talking about how well Apple develops its highly skilled workers, as well as the efforts made by Apple to develop technology for a better environment. This gap indicates a potential oversight in the media narrative, fail-

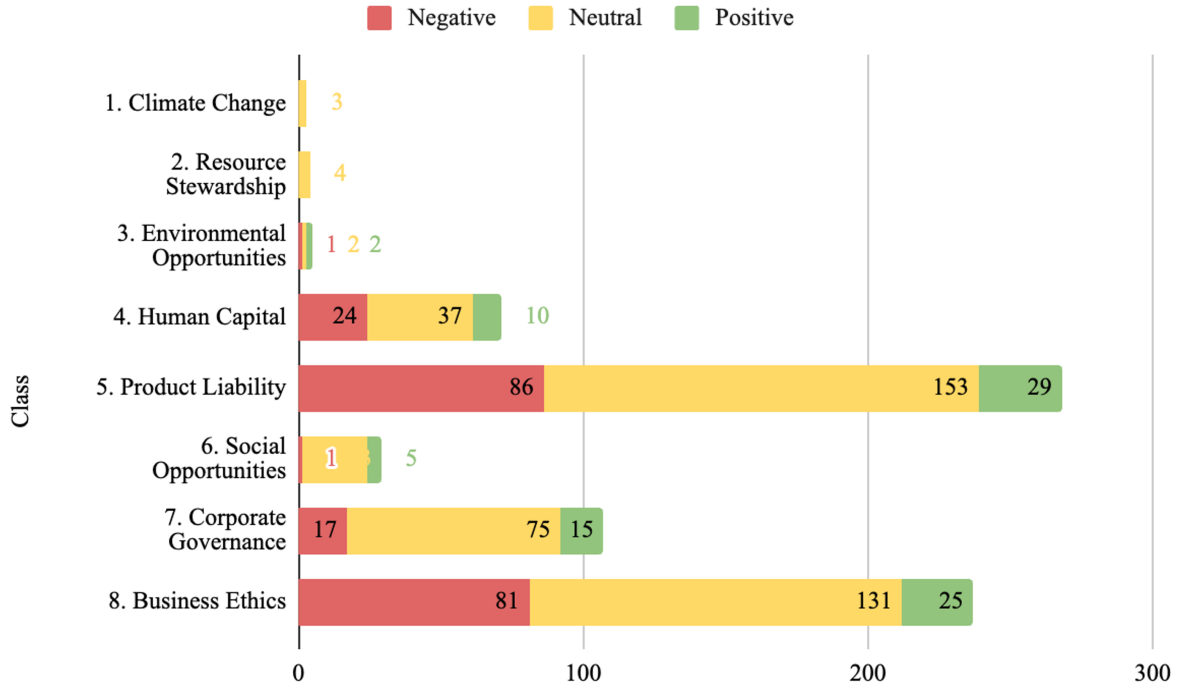


**Figure 6:** Apple's Sentiment Distribution Across Different ESG Categories

**Table 7:** Score Comparison Between Technology Companies

| | Apple | | Microsoft | | Alphabet | | Meta | |
|---|---|---|---|---|---|---|---|---|
| Categories | Score | Normalized Score | Score | Normalized Score | Score | Normalized Score | Score | Normalized Score |
| 1. Climate Change | 0 | 0.00 | 0 | 0.00 | 3 | 0.17 | 0 | 0.00 |
| 2. Resource Stewardship | 0 | 0.00 | 0 | 0.00 | 1 | 0.17 | 1 | 0.17 |
| 3. Environmental Opportunities | 1 | 0.20 | 1 | 1.00 | 2 | 0.09 | 2 | 0.67 |
| 4. Human Capital | -14 | -0.20 | -6 | -0.11 | -22 | -0.15 | -13 | -0.21 |
| 5. Product Liability | -57 | -0.21 | -44 | -0.26 | -130 | -0.28 | -136 | -0.29 |
| 6. Social Opportunities | 4 | 0.14 | 14 | 0.24 | 15 | 0.13 | -2 | -0.04 |
| 7. Corporate Governance | -2 | -0.02 | -9 | -0.05 | -4 | -0.02 | -20 | -0.13 |
| 8. Business Ethics | -56 | -0.24 | -97 | -0.37 | -186 | -0.34 | -128 | -0.30 |

ing to capture the complete picture of Apple's initiatives and its impact on sustainable business practices and responsible resource management.

#### 4.2.2. Microsoft

Microsoft stands as a global technological giant, with its main office situated in Washington, United States (Microsoft, n.d.-a). In terms of market capitalization, it is ranked as the third largest company worldwide, following Apple and Saudi Aramco, boasting a value of US$2.1 trillion as of 2022 (Johnston, 2022). Bill Gates and Paul Allen founded the company in 1975, initially focusing on the development and licensing of computer software, with their renowned operating system, Microsoft Windows, being a standout product (Zachary & Hall, 2023). Over the years, the company has expanded into several other product categories, including servers, productivity software, personal computers, consumer electronics, online advertising, and numerous other services (Zachary & Hall, 2023). The company has gained widespread recognition for its Windows operating system and Office Suite software. As of June 2023, Microsoft is a workplace for roughly 221,000 full-time employees around the globe (Microsoft, 2023). Currently, the company is under the leadership of Satya Nadella, who has been serving as the chairman and CEO since 2014 (Microsoft, n.d.-b).

There are marginally more articles from the New York Times discussing Microsoft's ESG initiatives compared to Apple, with 733 pieces. Of all the articles, the greatest focus was placed on 'Business Ethics,' which was the subject of 265 articles. The topic of 'Corporate Governance' also received considerable coverage, tallying up to 180 articles. This underscores the media's significant interest in the company's governance front, such as ethical considerations and leadership decisions. On the social facet, similar to Apple, 'Prod-

uct Liability' topics became the most discussed area with 167 articles. 'Social Opportunities' and 'Human Capital' also featured prominently, with 58 and 55 articles, respectively. The coverage on the environmental aspect was still low, with a cumulative count of only eight articles, hinting at a potential underreporting of Microsoft's environmental endeavors similar to what was observed with Apple's coverage.

Analogous to the sentiment of Apple's news, it was evident that the neutrality sentiment prevails in the articles about Microsoft's ESG practices, with the majority of articles—433 out of 733—displaying a neutral stance. The media kept reporting the news without a strong prejudice toward the company. Despite the neutral tendency, it is worth noting that the discussion around 'Business Ethics' was significantly critical, with 113 articles classified as negative compared to 136 neutral and only 16 positive ones, signaling strong media scrutiny of Microsoft's ethical conduct. Nevertheless, for 'Corporate Governance' and 'Human Capital,' it could be observed that over half of the articles maintained this neutral tone, with 125 and 35 articles, respectively. Regarding the 'Product Liability' angle, the number of negative articles was moderately prominent, accounting for just above 35% with 59 articles, while 93 remained neutral out of 167 articles. The environmental categories—namely 'Climate Change,' 'Resource Stewardship,' and 'Environmental Opportunities'—show a mixed but limited sentiment profile, with a small number of articles suggesting these topics have not been as contentious in the media. The sentiment for each category is depicted in Figure 7.

Analyzing Microsoft's performance through sentiment scores unveils compelling insights. As discovered previously, 'Business Ethics' registered as the most concerning area, more prevalent than those of Apple, with a total (normalized) score of -97 (-0.37). The score indicates that Microsoft had worse ethical concerns than its peers. Another area full of criticisms was 'Product Liability,' where Microsoft got a total (normalized) score of -44 (-0.26), slightly worse than Apple. Comparable to Apple's 'Corporate Governance' score, Microsoft's score was also close to neutral, with a normalized (total) score of -0.05 (-9). Microsoft tended to be better in terms of 'Human Capital' than Apple, albeit still inclined to negative nuanced with a normalized (total) score of -0.11 (-6). Microsoft attained a much better score in terms of 'Social Opportunities' than Apple, with a normalized (total) score of 0.24 (14), indicating more societal actions from Microsoft got more attention from the media. Regrettably, on the environmental front, there were only a few news pieces available to be generalized. The detailed scoring is available in the Appendix.

Delving into the 'Business Ethics' news, Microsoft was often being criticized regarding antitrust and abuse of power matters (e.g., Lohr, 2020; Markoff, 2003; Weise and McCabe, 2022). A significant portion of this coverage stems from an earlier era, with 81 out of 113 articles concentrated between 2003 and 2009. In contrast, the subsequent 13 years accounted for just 32 articles, suggesting a potential shift toward more ethical operations by Microsoft, though criticisms persist. Concerning the 'Product Liability' subject, security vulnerabilities, particularly breaches in Microsoft's email systems, have been a focal point of media attention (e.g., Conger and Frenkel, 2021; Scott, 2015). Challenges in 'Human Capital' have also emerged, linked to immigration issues and job cuts (e.g., Frenkel, 2018; Wingfield, 2016a). Nonetheless, Microsoft has made notable efforts to contribute positively to society, including initiatives to support affordable housing and to enhance the accessibility of essential services for public servants (e.g., Vance, 2010; Weise, 2019).

MSCI places Microsoft within the software and services industry and grants it the highest rating (AAA) among 484 companies (MSCI, 2023a). These prestigious rankings originate from MSCI's assessment of Microsoft's exemplary performance in critical areas such as human capital development, safeguarding privacy and data, opportunities in clean tech, and mitigating carbon emissions (MSCI, 2023a). However, the results from analyzing the news articles show the investment in Microsoft's talented employees seems to be underrepresented. Interestingly, in terms of privacy & data security, MSCI evaluates Microsoft differently in regard to controversies from the media[10]. Microsoft is considered to be involved in severe-to-moderate level controversies—similar to what I found in the news—while MSCI still rates Microsoft as the leader in this aspect compared to its competitors (MSCI, 2023a). Additionally, the media's limited coverage of Microsoft's environmental attempts, with merely a single article on both its cleantech initiatives and carbon reduction strategies, raises questions about the adequacy of information to fully evaluate the company's environmental performance.

In a similar marking to Apple, MSCI assigns Microsoft a less favorable rating on the topic of business ethics and considers it to be average in corporate governance (MSCI, 2023a). Reciprocal with this rating, the sentiments found in the media reports suggest that Microsoft could benefit from a reassessment of its ethical procedures to address concerns about its dominant market position and abuse of power. This is also in line with the controversies indicator by MSCI, in which Microsoft is considered in moderate to severe controversies. As for corporate governance, the majority of media sentiments remain neutral but tilt slightly towards the negative, broadly matching MSCI's findings. This indicates that Microsoft's approach to governance—encompassing both its control structure and ethical conduct—requires further scrutiny and potential enhancement to uphold and improve its standing in these critical aspects.

#### 4.2.3. Alphabet Inc. (Google)

Alphabet Inc. (Alphabet) is a conglomerate holding company of Google, a multinational technology company that provides a diverse range of products and services, including advertising, operating systems, hardware, web browsers,

[10] MSCI ESG Controversies serves as an additional indicator to assess companies' profiles, focusing on their actual or alleged involvement in activities that have adverse impacts, as reported by the media.

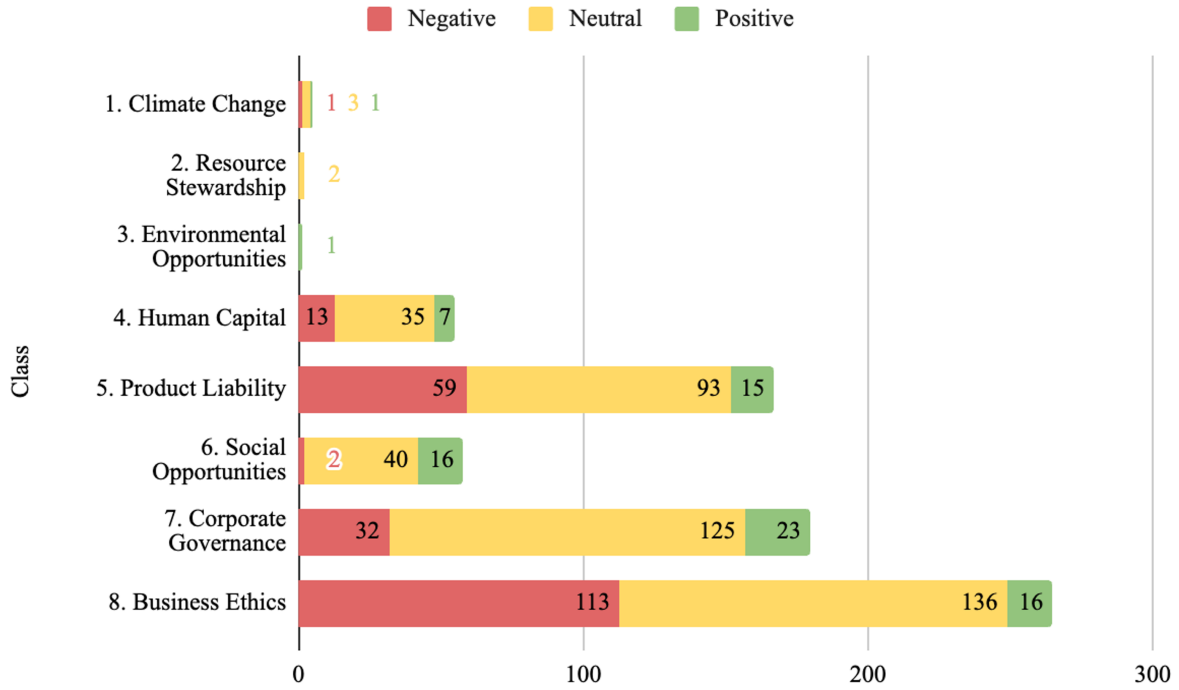


**Figure 7:** Microsoft's Sentiment Distribution Across Different ESG Categories

and cloud services (Alphabet Inc., 2022). Its market capitalization ranked fourth worldwide, with its worth recorded at US$1.54 trillion in 2022 (Johnston, 2022). Google, founded by Sergey Brin and Larry Page in 1998, initially started as a search engine and has since expanded to offer over 50 different internet services and products (Hall & Hosch, 2023). Some of its most notable offerings include Google Chrome, Gmail, Android, and Google Maps. In a significant move in 2015, Google underwent a reorganization, leading to the formation of Alphabet as a new public holding company (Alphabet Inc., 2015). The company's primary headquarters, known as the Googleplex, is situated in Mountain View, California, United States, and is the workplace for approximately 190,000 employees (Alphabet Inc., 2022). In 2019, Sundar Pichai took on the role of CEO for both Google and Alphabet, succeeding co-founders Larry Page and Sergey Brin (Alphabet Inc., 2019). Since Google is the biggest company under Alphabet, the discussion below mainly contains Google's news articles.

The New York Times has provided more extensive coverage of Alphabet and Google's ESG initiatives than it has for Apple and Microsoft. Out of 1,506 pieces spanning eight topics, 'Business Ethics' predominated, being the subject of 551 articles, which is over one-third of the total. The theme of 'Product Liability' also stood out with a significant count of 472 articles, over 30% of all articles. Meanwhile, 'Corporate Governance' and 'Human Capital' also received considerable attention, with 182 and 142 articles, respectively. The ranking of the most reported topics for Alphabet/Google closely mirrored that of Microsoft, with 'Business Ethics,' 'Product Liability,' and 'Corporate Governance' receiving the most spotlight. Social impact was also a notable subject, with 'Social Opportunities' discussed in 113 articles. The environmental category received more attention compared to their industry counterparts, with 46 articles in total, although it was still less compared to other subjects. In this spectrum, 'Environmental Opportunities' led with 22 articles, followed by 'Climate Change' with 18, and 'Resource Stewardship' with 6 articles.

The sentiment analysis of The New York Times articles on Google's ESG efforts reveals a predominately neutrality across the topics, with 933 out of 1506 articles maintaining an unbiased tone. However, in the domain of 'Business Ethics,' a critical perspective is noticeable, with 228 articles carrying a negative sentiment, almost outstripping the 281 neutral articles. This might suggest a keen media vigilance on Google's ethical practices. In contrast, 'Human Capital' and 'Corporate Governance' topics show a preference for neutrality, with 98 and 136 articles, respectively, indicating less contentious coverage. 'Product Liability' received quite controversial coverage, with a substantial 154 articles skewed towards a negative sentiment, highlighting concerns or issues in this area, although 294 articles remain neutral. 'Social Opportunities' presents a more positive outlook with 20 positive articles. Environmental issues showcase a diverse but generally limited sentiment range. 'Climate Change' has 13 neutral and 4 positive articles—only one article is considered negative. This could imply that while these topics are covered, they do not ignite as much as other topics in the media. The breakdown of sentiment across each category is illustrated in Figure 8.

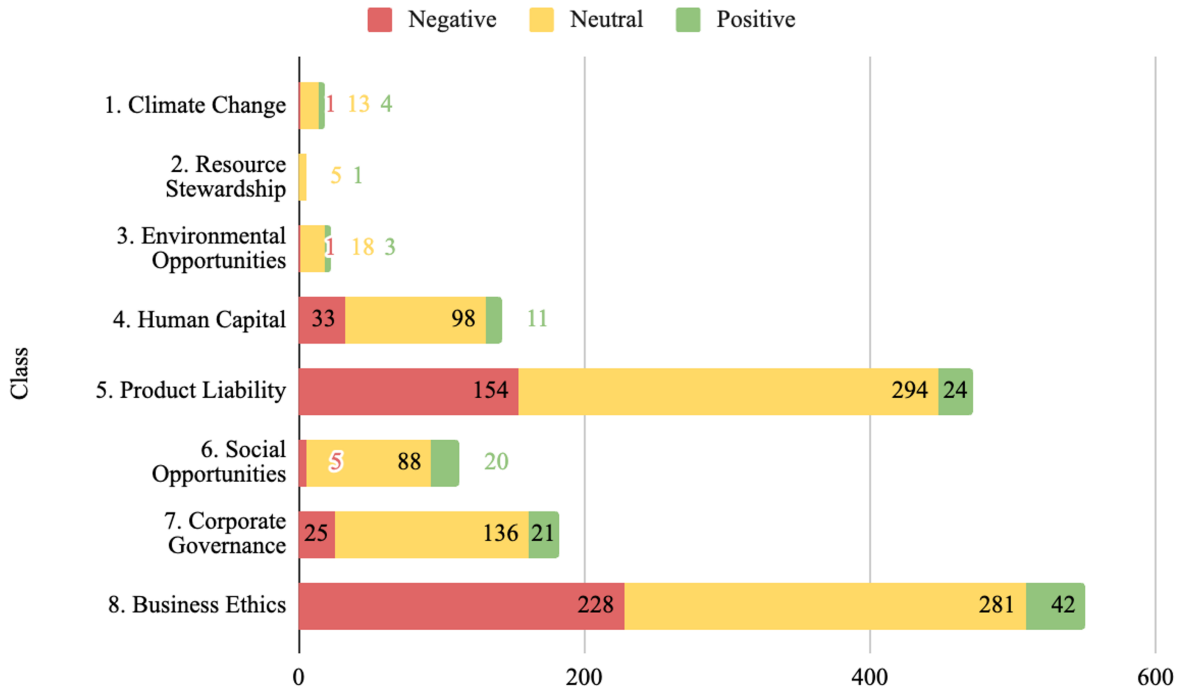


**Figure 8:** Alphabet's Sentiment Distribution Across Different ESG Categories

The sentiment assessment for Google's ESG performance presents a varied impression. In contrast to Apple and Microsoft, scoring the environmental features of Google's performance is more straightforward. This clarity comes from a broader range of sentiments represented across the categories. Specifically, Google demonstrates a positive tilt in this sector, with the areas of 'Climate Change,' 'Resource Stewardship,' and 'Environmental Opportunities' achieving normalized (total) scores of 0.17 (3), 0.17 (1), and 0.09 (2), respectively. These figures suggest a favorable perception of Google's efforts in the environmental matter. Conversely, 'Human Capital' reflects a notable drop with a normalized (total) score of -0.15 (-22), indicating areas of concern within this sphere. In the other social contexts, the 'Product Liability' category is markedly the most critical, with a significant negative normalized (total) score of -0.28 (-130), while 'Social Opportunities' conversely portrays a brighter spot with a positive normalized (total) score of 0.13 (15). The 'Product Liability' aspect is even more concerning than those of Microsoft and Apple, with a normalized score of -0.26 and -0.21, respectively. In terms of 'Social Opportunities,' Microsoft is also still leading with a normalized score of 0.24. The detailed scoring is available in the Appendix.

Moving to the governance domain, Google's 'Corporate Governance' presents nearly balanced sentiments with a thin negative normalized (total) score of -0.02 (-4), which aligns it closely with its industry counterparts. However, 'Business Ethics' emerges as the most concerning area, with a stark normalized (total) score of -0.34 (-186), signifying strong negative sentiment and possibly considerable criticism in this area. Although Google's normalized score in 'Business Ethics' is somewhat better than Microsoft's (-0.37), it is important to note that Google's total negative score is -186, compared to Microsoft's -97. This difference may point to Google facing more intensive media scrutiny over its ethical practices.

Reflecting on the media coverage of Google in the environmental landscape, there is a noticeable acknowledgment of the company's investments in renewable energy and other sustainable technologies. Several reports, although categorized as neutral, highlight Google's promising endeavors to address climate change (e.g., Austen, 2019; Hardy, 2016). In the social scene, particularly in 'Product Liability,' various issues have surfaced, such as problematic user data tracking, the misuse of products for extortion, and defects in product quality (e.g., Hill and McCabe, 2022; Morales, 2022; Woo, 2022). In the area of 'Human Capital,' Google has faced criticism regarding its treatment of employees, particularly temporary and contract workers (e.g., Scheiber, 2020; Wakabayashi, 2021). Nevertheless, the company has also gained recognition for its initiatives aimed at enhancing quality of life. An example of such efforts includes the development of healthcare tools designed to support doctors in making informed decisions (e.g., Grady, 2020). On the most concerning side, 'Business Ethics,' the media has intensely scrutinized Google's ethical practices, especially in relation to antitrust issues. Google's involvement in various lawsuits across different regions and for different reasons has drawn considerable attention from the news outlet (e.g., Satariano, 2022a; Tracy, 2021).

As stated by MSCI, Alphabet has been assigned a 'BBB' rating, positioning it as average within the interactive media & services sector, which comprises 65 companies (MSCI,

2023a). The assessment highlights several concerns, such as corporate conduct, governance practices, employee development, and the pursuit of clean technology innovations (MSCI, 2023a). Corresponding to the findings from the media, the company is heavily criticized for its ethical behaviors. In the sphere of corporate governance, Alphabet appears to maintain a neutral stance, comparable to its peers like Microsoft and Apple, though MSCI categorizes it as a laggard (MSCI, 2023a). Points of disagreement also emerge in the opportunities in clean tech, where MSCI evaluates the company as lagging behind (MSCI, 2023a), whereas news reports indicate the company has launched multiple initiatives in this field. Nonetheless, there is a consensus between this research and MSCI's findings that Alphabet has taken steps towards reducing its carbon footprint (MSCI, 2023a).

#### 4.2.4. Meta Platforms, Inc.

Meta Platforms, Inc. (Meta), previously recognized as Facebook, is a major technology corporation based in Menlo Park, California (Meta Platforms Inc., 2022). The company owns and operates a variety of products and services, including Facebook, Instagram, Messenger, and WhatsApp (Meta Platforms Inc., 2022). Currently, the technology giant is ranked 10th in terms of market capitalization in the world, with a value of US$449 billion (Johnston, 2022). Since Mark Zuckerberg founded Facebook in 2004 along with Dustin Moskovits, Chris Hughes, and Eduardo Saverin, Meta has grown from a social media platform connecting friends and family to pioneering immersive experiences through augmented and virtual reality (Meta Platforms Inc., n.d., 2022). Zuckerberg continues to play a pivotal role, leading the company as both chairman and CEO (Meta Platforms Inc., n.d.). As of the end of December 2022, the company recorded a total of 86,482 employees, inclusive of the roughly 11,000 workers affected by the massive layoff announced in November of the same year (Meta Platforms Inc., 2022).

The New York Times' coverage of Meta's ESG efforts encompasses a total of 1,161 articles across the topics, more than Apple and Microsoft's coverage. Similar to Apple, the most prevalent subject in this collection is 'Product Liability,' which dominates the spotlight with 470 articles, making up a substantial portion of the total coverage. This is followed by 'Business Ethics,' which is the focus of 422 pieces. 'Corporate Governance' ranks third with 150 articles, similar order to its technology peers. On the social front, 'Human Capital' and 'Social Opportunities' are notable topics, covered in 61 and 47 articles, respectively. Environmental topics remain underrepresented, with a total of just 11 articles; only Alphabet receives more coverage in the technology sector, featuring in over 20 articles. In the environmental domain, 'Resource Stewardship' interestingly is the subject of a greater number of articles, with 6 pieces, compared to 'Climate Change' and 'Environmental Opportunities,' which are covered in 2 and 3 articles, respectively.

The sentiment analysis of Meta's performance shows a pattern consistent with the other three technology companies, where the impartial view dominates the results, accounting for over 60% of the articles. Yet, 'Product Liability' stands out with 160 negative mentions, marking it as the most criticized topic despite a generally neutral coverage. Similarly, the domain of 'Business Ethics' is also heavily critiqued, with 140 negative articles against 270 neutral ones. 'Business Ethics' is one of the most concerning topics for technology companies, not just Meta. The scrutiny suggests the media's close watch on the ethical practices of large tech firms. Coverage on 'Human Capital' and 'Corporate Governance' predominantly maintain a neutral stance, numbering 38 and 106, respectively, pointing to less controversial media reporting in these areas. 'Social Opportunities' and environmental topics display a mix of sentiments but are covered less extensively. Figure 9 depicts the sentiment distribution across different categories.

Meta's sentiment scores highlight areas of concern, particularly within the social and governance categories. While environmental aspects like 'Resource Stewardship,' 'Environmental Opportunities,' and 'Climate Change' exhibit neutral to positive scores, coverage in these areas is sparse. In contrast, 'Product Liability' and 'Business Ethics' face notable scrutiny, with significant negative normalized (total) scores of -0.29 (-136) and -0.30 (-128), respectively. These scores are almost identical to those of Alphabet, indicating that these two tech giants are closely watched by the media regarding their products and ethical practices. 'Human Capital' also registers a substantial negative normalized (total) score of -0.21 (-13), marking the lowest performance among the other technology companies. In the typically more positive 'Social Opportunities' category, Meta surprisingly scores slightly negative at -0.04 (-2) normalized (total) score. This negative sentiment extends to 'Corporate Governance,' where Meta's normalized (total) score of -0.13 (-20) is the least favorable when compared with its sector peers. The detailed scoring is available in the Appendix.

Upon examining the articles in detail, it becomes apparent that the 'Business Ethics' theme is largely characterized by improper business practices and antitrust litigation (e.g., C. Kang et al., 2019; Satariano, 2020). Moving to the other governance facet, Meta received criticisms for the decision-making and conduct of its leadership, as well as the departure of key executives (e.g., Isaac, 2018; Satariano and Frenkel, 2022). In the 'Product Liability' category, which garners the most negative attention for Meta, several issues are highlighted by the media, including privacy breaches, potential for dangerous behaviors, and unreliable services (e.g., Isaac, 2021; Mac, 2022; Satariano, 2022b). Common challenges like antitrust and privacy issues are also faced by Meta's industry peers. Nonetheless, given the significant role Meta's products play in social interactions, there is heightened media concern over the ways these products can harm individuals, and it is suggested that the company has not adequately addressed these risks.

MSCI grades Meta with the lowest rating, CCC, compared to the other companies within the interactive media & services industry, the same sector as Alphabet (MSCI, 2023a). Meta is particularly lagging in areas of ethical conduct and

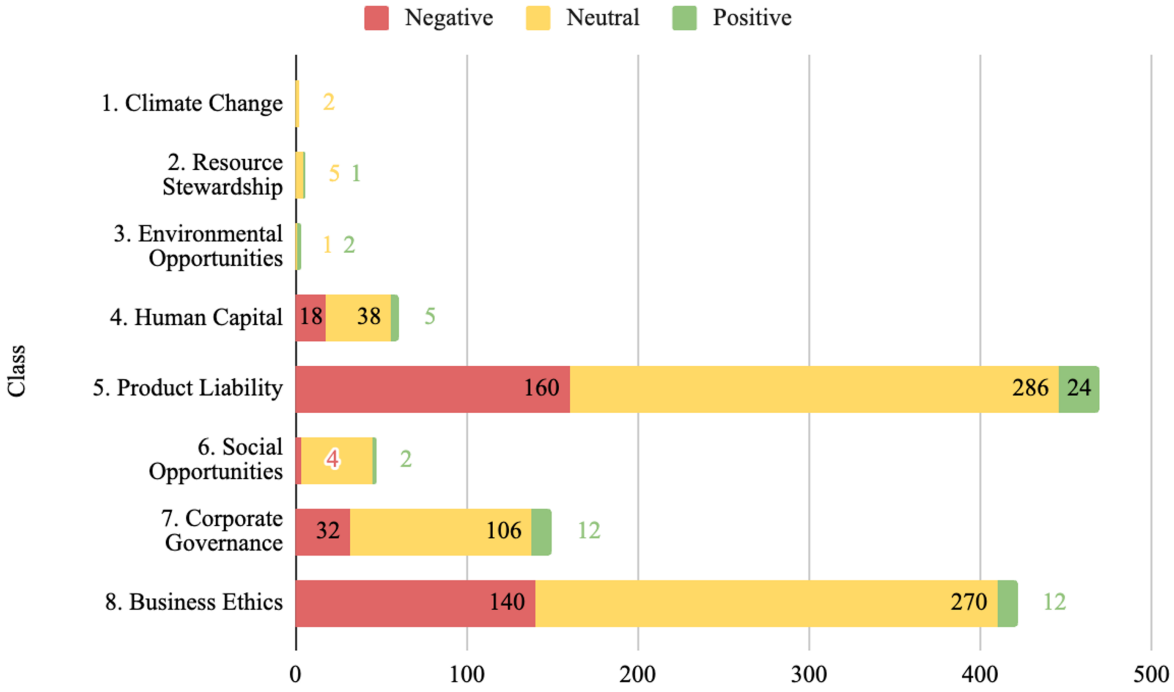


**Figure 9:** Meta Platforms' Sentiment Distribution Across Different ESG Categories

human capital development (MSCI, 2023a). These concerns echo the media's criticisms of Meta's business practices and allegations of anti-competitive actions. The media reports also disclose Meta's plans to pause hiring and reduce staff (e.g., Frenkel, 2022a, 2022b). MSCI's controversy indicator supports these points, suggesting Meta is entangled in moderate to severe labor rights disputes (MSCI, 2023a). However, when it comes to corporate governance and privacy, MSCI considers Meta's performance average, despite recognizing the company's involvement in moderate to severe governance disputes and serious privacy & data security issues (MSCI, 2023a). These MSCI assessments of controversy indicators are consistent with the results from this study, which indicate negative scores for Meta in these two categories. The sole area where MSCI acknowledges Meta's leadership is in its efforts to slash carbon emissions (MSCI, 2023a).

#### 4.2.5. The Coca-Cola Company

Started in 1886 when Dr. John Pemberton sold his newly crafted syrup at Jacob's Pharmacy, laying the foundation for what would become a multinational total beverage corporation (The Coca-Cola Company, n.d.-b). Today, the company has grown far beyond its famous Coca-Cola drink, managing an array of brands like Sprite and Fanta and extending its reach to over 200 countries and territories (The Coca-Cola Company, n.d.-b). The company also offers a diverse product lineup that includes coffee, tea, juice, value-added dairy, plant-based beverages, and innovative new drinks (The Coca-Cola Company, 2022). At the helm is James Quincey, who has been with the company since 1996 and currently oversees its operations and 82,500 employees as the chairman and CEO (The Coca-Cola Company, n.d.-a, 2022). Under his leadership, this Atlanta-based company has continued to thrive, with a reported market value of US$242 billion (Britannica, 2023a).

The examination of Coca-Cola's ESG initiatives is covered in a total of 101 articles, a smaller number compared to the extensive coverage of major technology firms. Among these, 'Corporate Governance' emerges as the most discussed topic with 35 articles. 'Business Ethics' is another significant topic, featuring in 20 articles. The environment-related categories show varied coverage, with 'Resource Stewardship' leading with 13 articles. On social issues, the coverage for each topic is distributed relatively evenly. These articles are dominated by neutral sentiment, with a total of 74 out of 101 articles. 'Corporate Governance' sees an equal distribution of positive and negative sentiment, each with four articles. 'Business Ethics' faces the most scrutiny, reflected in six articles with a negative sentiment. Regarding environmental issues, 'Resource Stewardship' draws the most concern with five negative articles. All topics in the environmental sphere do not garner positive highlights from the media, indicating challenges in this area. Similarly, social topics also do not receive positive media attention. Instead, the coverage is primarily neutral, though some negative perspectives are present. Figure 10 provides a visual breakdown of the sentiments across these categories.

As reflected by sentiment scores, Coca-Cola's performance indicates several areas of concern across ESG topics. 'Resource Stewardship' incurs the most negative sentiment, with a normalized (total) score of -0.38 (-5). However, when taking a deeper look into the articles, there are some mis-

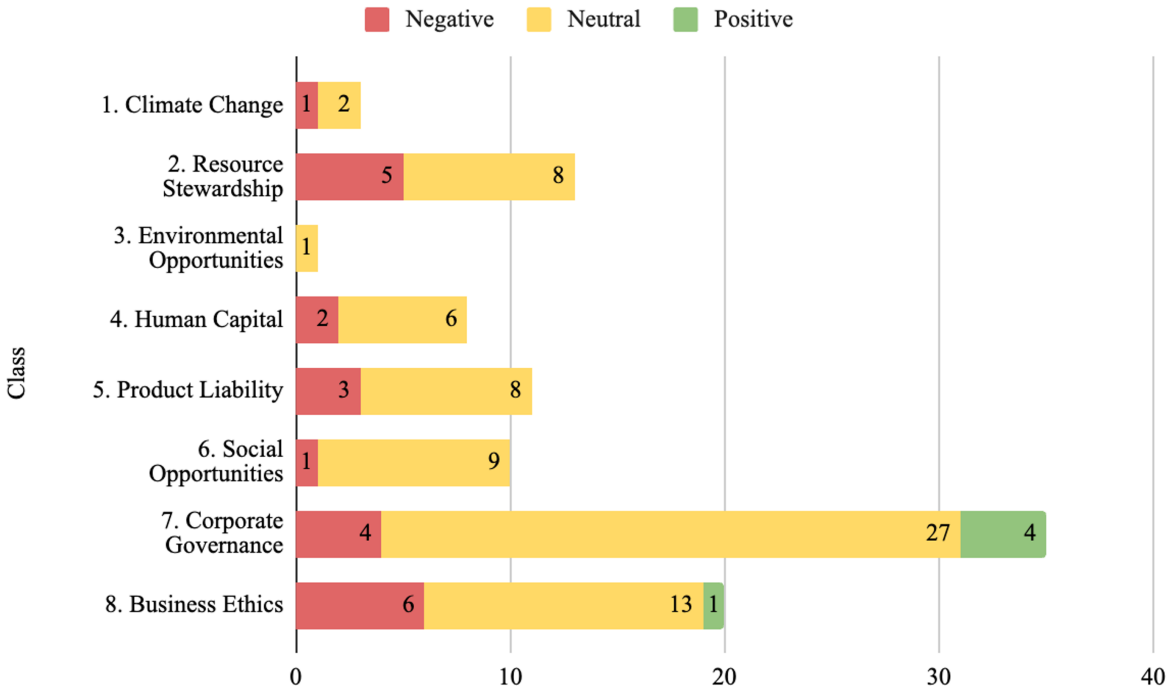


**Figure 10:** The Coca-Cola Company's Sentiment Distribution Across Different ESG Categories

classifications in terms of the sentiment, where the articles are considered negative and neutral, whereas they discuss positive efforts by Coca-Cola, such as recycling and water replenishment initiatives (e.g., Corkery, 2019; Schwartz, 2015). This means the scoring might not reflect the actual performance of the company in this domain. 'Business Ethics' also draws a notably negative view, sharing a normalized score of -0.25 with 'Human Capital.' Coca-Cola faces some critiques from the media for its ethical conduct, such as the alleged drug-smuggling operation and insider trading (e.g., Lattman, 2012; Mele, 2016). Several news pieces also covered the company's action to cut jobs (e.g., Strom, 2015). In another social theme, Coca-Cola scores -3 in total score and -0.27 in normalized score in 'Product Liability,' pointing to perceived issues in its products and logistics. Particularly, the media criticizes how Coca-Cola often sugarcoats the calories or nutrition facts, adds questionable ingredients to the products, and its supply chain management (e.g., Board, 2015; Southall, 2015; Strom, 2014a). 'Corporate Governance,' although it breaks even with a score of 0, the company still receives critiques about pay for its executives (e.g., Eavis, 2014). Table 8 exhibits the scoring comparison between the beverage and the pharmaceutical companies. The detailed scoring is available in the Appendix.

Rated as AAA, the highest rating in MSCI ESG Ratings, Coca-Cola stands out as the top performer among a hundred companies in the beverage sector (MSCI, 2023a). The company excels in several critical areas, such as management practices, developing healthful products, conserving water, ensuring the safety and quality of products, safeguarding employee health, and reducing the carbon emissions of its products (MSCI, 2023a). However, Coca-Cola's performance in ethical business conduct and waste management from packaging is seen as less impressive, with MSCI calling it mediocre (MSCI, 2023a). Sentiment analysis also casts a shadow on Coca-Cola's reputation, as ethical concerns arise from insider trading scandals. MSCI's data on controversies supports this view, indicating that Coca-Cola has been involved in a moderate amount of bribery and fraud cases (MSCI, 2023a).

Regarding 'Resource Stewardship,' MSCI views Coca-Cola's handling of packaging materials and waste as typical (MSCI, 2023a), but news outlets have noted the company's efforts to recycle. Nevertheless, both the media and MSCI acknowledge the company's efforts in its water management practices (MSCI, 2023a), such as the effort to replenish water that it uses around the globe (e.g., Schwartz, 2015). However, there's a difference of opinion on 'Product Liability.' MSCI ranks Coca-Cola highly in this category, suggesting leadership and responsibility in product nutrition (MSCI, 2023a). In contrast, sentiment analysis suggests Coca-Cola faces challenges here, with concerns over ingredients and nutritional information that may pose health risks to customers (e.g., Board, 2015).

#### 4.2.6. PepsiCo

PepsiCo, Inc. was formed in 1965 through the merger of Pepsi-Cola Company and Frito-Lay, Inc (Britannica, 2023b). The original Pepsi-Cola, crafted by Caleb D. Bradham in 1898, gained formal corporate status in 1919 (Britannica, 2023b; PepsiCo Inc., 2022). Nowadays, PepsiCo is recognized as a prominent force in both the beverages and convenience foods sector, home to well-known brands such as Pepsi, Lay's, Cheetos, and Quaker (PepsiCo Inc., 2022).

**Table 8:** Score Comparison Between Beverage and Pharmaceutical Companies

| | Coca-Cola | | PepsiCo | | Pfizer | | Johnson & Johnson | |
|---|---|---|---|---|---|---|---|---|
| Categories | Score | Normalized Score | Score | Normalized Score | Score | Normalized Score | Score | Normalized Score |
| 1. Climate Change | -1 | -0.33 | 0 | 0.00 | N/A | N/A | N/A | N/A |
| 2. Resource Stewardship | -5 | -0.38 | -2 | -0.67 | 0 | 0.00 | -1 | -1.00 |
| 3. Environmental Opportunities | 0 | 0.00 | N/A | N/A | 0 | 0.00 | 1 | 1.00 |
| 4. Human Capital | -2 | -0.25 | 0 | 0.00 | -4 | -0.27 | -1 | -0.20 |
| 5. Product Liability | -3 | -0.27 | -2 | -0.29 | -19 | -0.19 | -49 | -0.40 |
| 6. Social Opportunities | -1 | -0.10 | 4 | 0.40 | 43 | 0.30 | 11 | 0.23 |
| 7. Corporate Governance | 0 | 0.00 | 0 | 0.00 | -4 | -0.09 | -2 | -0.13 |
| 8. Business Ethics | -5 | -0.25 | -5 | -0.56 | -14 | -0.26 | -24 | -0.65 |

Similar to its industry peer, Coca-Cola, PepsiCo products are available in more than 200 countries and territories (PepsiCo Inc., 2022). The company operates out of its North Carolina base under the direction of Ramon L. Laguarta, who has been the CEO since 2018 and as Chairman since 2019 (PepsiCo Inc., n.d., 2022). With a global workforce of approximately 315,000 as of the end of 2022, PepsiCo continues to be a major player in the food and beverage industry (PepsiCo Inc., 2022).

Even less extensive than Coca-Cola's, PepsiCo's ESG efforts are only covered in 61 articles, with 'Corporate Governance' being the primary focus in 24 articles, indicating a significant media focus on the company's management practices and policies. 'Social Opportunities' is another area of interest, highlighted in 10 articles, showing the media's attention to Pepsi's efforts in social engagements and community impact. The topic of 'Business Ethics' is covered in 9 pieces, reflecting considerations of PepsiCo's ethical conduct. Environmental issues such as 'Climate Change' and 'Resource Stewardship' are less frequently discussed, with 5 and 3 articles, respectively, and there are no articles addressing 'Environmental Opportunities,' suggesting these areas may be less scrutinized or perhaps better managed. 'Human Capital' and 'Product Liability' categories are also observed, with 3 and 7 articles each. The distribution of the sentiment for PepsiCo is portrayed in Figure 11.

The sentiment analysis reflects a primarily neutral stance, with the majority of articles, 44 out of 61, not leaning towards either a positive or negative sentiment. 'Corporate Governance' is predominantly viewed in a neutral light, suggesting a balanced media perspective on the company's management practices. However, with over half of the publicity being negative, 'Business Ethics' becomes a topic of concern over PepsiCo's ethical conduct. 'Product Liability' is discussed neutrally in 5 articles, with 2 articles expressing a negative sentiment. 'Social Opportunities' stands out with a more favorable view, dominating the positive sentiments. Environmental categories receive limited but mostly neutral coverage, with 'Climate Change' not receiving any positive or negative sentiments, while 'Resource Stewardship' leans toward the negative.

PepsiCo's ESG performance sentiment scores highlight critiques in certain areas while also recognizing some positive actions. 'Resource Stewardship' and 'Business Ethics' have a negative sentiment with a normalized (total) score of -0.67 (-2) and -0.56 (-5), respectively. In contrast, 'Social Opportunities' stands out positively with a normalized (total) score of 0.4 (4), showcasing PepsiCo's commitment to social impact, such as increasing access to chickpeas in Ethiopia and awarding grants to businesses promoting social causes (e.g., Newman, 2011; Strom, 2011). However, PepsiCo's commitment to reducing sugary drink calories (e.g., Strom, 2014b), indicating proactive steps in 'Social Opportunities,' is classified as neutral instead of positive. Many articles under 'Corporate Governance' are regarded as neutral, such as those related to leadership changes and investor relations (e.g., Creswell, 2018; Picker, 2016). The deceptive marketing and infiltration of political agenda in its charity organization contributes to its 'Business Ethics' negative assessment (e.g., Strom, 2010; Vega, 2011). In the 'Product Liability' articles, the media highlights the risk of type 2 diabetes for countries with higher usage of high-fructose corn syrup, like in PepsiCo's products (e.g., Bittman, 2012). The detailed scoring is available in the Appendix.

In comparison with other companies in its sector, PepsiCo is ranked AA by MSCI, which is slightly lower than Coca-Cola, yet PepsiCo is still seen as a top company (MSCI, 2023a). The issue of how PepsiCo gets its raw materials is noted by MSCI as a concern, though it is not widely discussed in the news (MSCI, 2023a). MSCI considers PepsiCo's performance in areas like company management, business ethics, waste management, and worker health and safety to be average (MSCI, 2023a). This matches with sentiment analysis findings, which suggest that PepsiCo's approach to managing its business is balanced. Nonetheless, the sentiment analysis pointed out some problems with PepsiCo's moral practices, whilst MSCI rated these as typical for the industry. MSCI views PepsiCo as leading the way in promoting better nutrition and health, managing water use, and ensuring product safety and quality, similar to Coca-Cola (MSCI, 2023a). However, the sentiment analysis does not fully acknowledge PepsiCo's improvements in these areas; instead, it highlights potential health risks from its product ingredients.

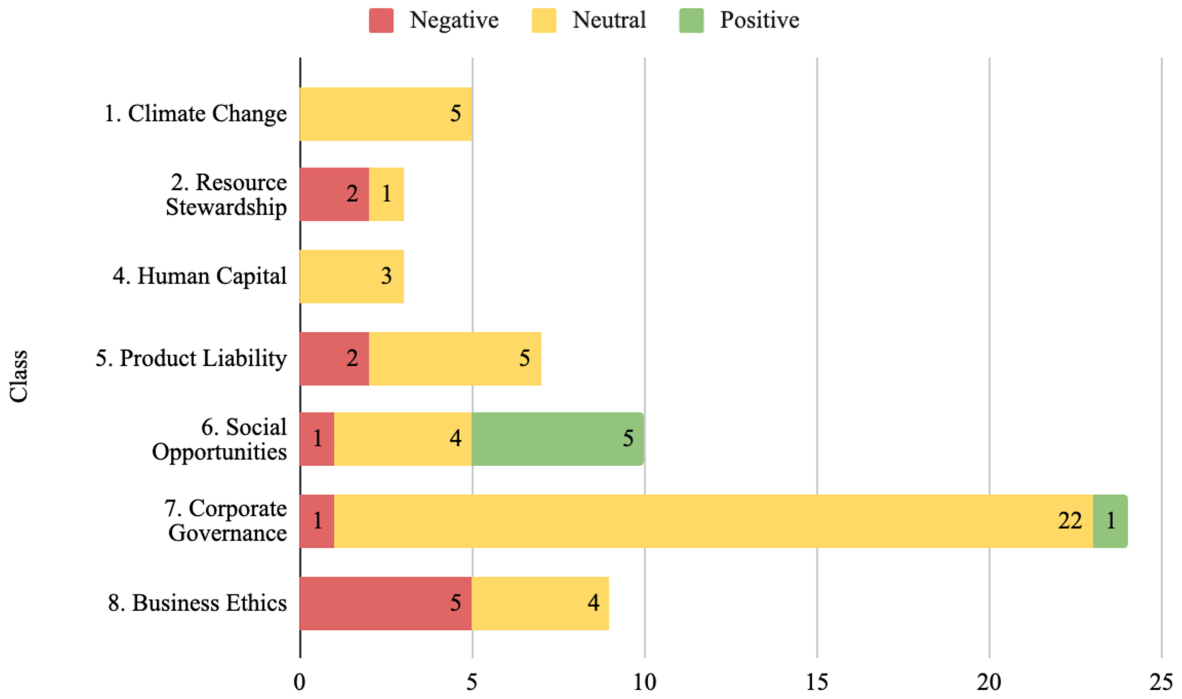


**Figure 11:** PepsiCo's Sentiment Distribution Across Different ESG Categories

4.2.7. Pfizer, Inc.

Pfizer, Inc. (Pfizer) is one of the biggest research-based pharmaceutical and biomedical corporations in the world, engaged in the creation, production, and global distribution of medical treatments (Nolen, 2023). Founded over a century ago in 1849 by the German chemist and entrepreneur Charles Pfizer and his cousin, Charles Erhart, the company has made significant strides in the field of medicine (Nolen, 2023). Notably, during the recent COVID-19 pandemic, Pfizer was at the forefront, developing the Comirnaty vaccine in collaboration with BioNTech, signifying an achievement in the global fight against the virus (Pfizer Inc., 2022). Pfizer is also recognized for pioneering medications such as Lipitor and Viagra, extending its reach to over 185 countries and territories (Pfizer Inc., 2022). The company is currently led by Dr. Albert Bourla as the chairman and CEO, guiding a dedicated workforce of 83,000 employees based out of their headquarters in New York (Pfizer Inc., n.d., 2022).

Pfizer receives extensive coverage of its ESG initiatives, spanning 364 articles. Among these, 'Social Opportunities' is the most discussed, with 143 articles highlighting the company's efforts in this area. 'Product Liability' also receives substantial attention, with 102 articles addressing the impacts of Pfizer's products. 'Business Ethics' is another focus, with 54 articles discussing the company's ethical practices. Following closely behind, 'Corporate Governance' is examined in 47 articles, reflecting scrutiny of the company's leadership and management decisions. 'Human Capital' emerges as well, with 15 articles potentially focusing on labor practices and workforce management. Environmental themes are less prominent; 'Resource Stewardship' and 'Environmental Opportunities' only account for 3 articles in total, pointing to a less pronounced focus on these issues. In addition, there is a noticeable absence of coverage of 'Climate Change,' indicating either a lack of effort or reporting in this critical area of ESG. Figure 12 visualizes the sentiment distribution for each category of Pfizer.

The sentiment analysis of Pfizer's ESG efforts skews towards a neutral viewpoint, with 204 articles maintaining an impartial tone. The 'Product Liability' discourse is mixed, with 43 articles reflecting negative sentiment and 24 articles expressing positivity, pointing to varied public reception and media reporting on the impact of Pfizer's products. 'Social Opportunities' reaps the brightest sentiment, with 49 articles acknowledging positive actions, significantly outweighing the 6 negative pieces, while environmental discussions under 'Resource Stewardship' and 'Environmental Opportunities' are discussed minimally but largely neutral. The majority of articles under 'Corporate Governance' depict a neutral stance, while 'Business Ethics' emerges as an area of concern, with 16 articles bearing a negative sentiment, possibly highlighting scrutiny over the company's ethical conduct.

It is evident that Pfizer's 'Social Opportunities' stands out with a positive total (normalized) score of 43 (0.30), suggesting a strong positive reception in this angle. However, areas such as 'Product Liability' and 'Business Ethics' are notable points of concern. 'Product Liability' has a particularly low total (normalized) score of -19 (-0.19), while 'Business Ethics' shows significant room for improvement with a total (normalized) score of -14 (-0.26). Concerns are also apparent in the 'Human Capital' category, where the negative sentiment is more pronounced at a normalized score of -0.27,

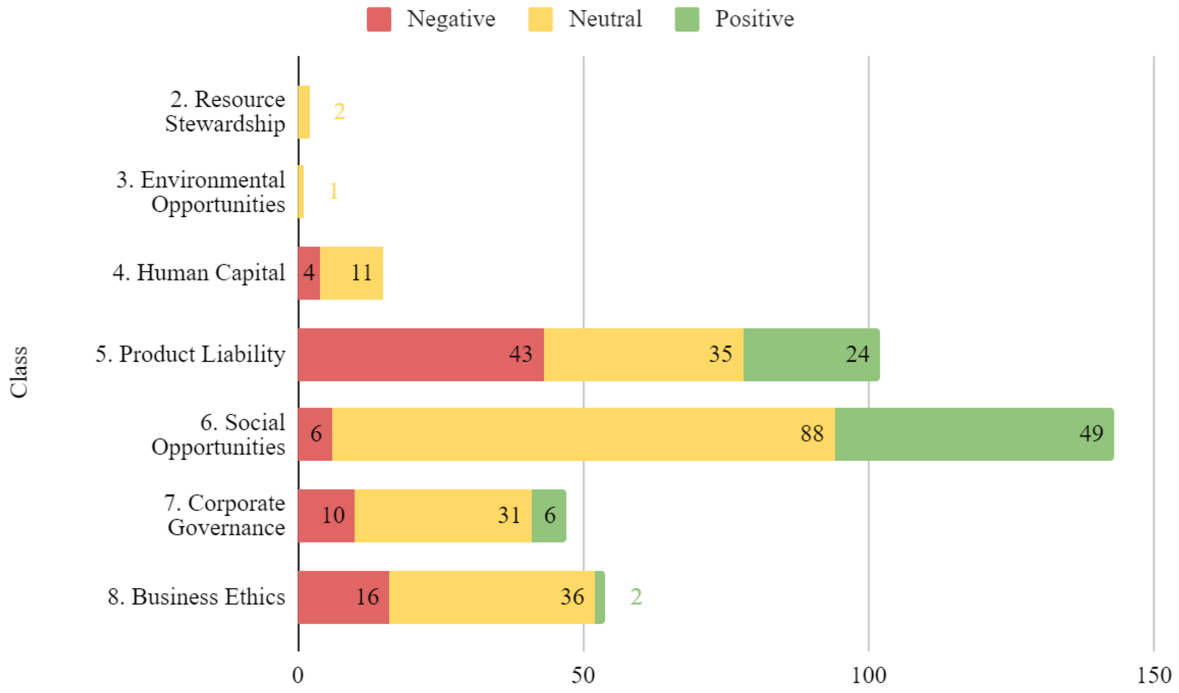


**Figure 12:** Pfizer's Sentiment Distribution Across Different ESG Categories

despite a total score of just -4. 'Corporate Governance' parallels this with a total score of -4 but with a less negative normalized score of -0.09, suggesting a predominantly neutral perception. In contrast, environmental categories have neither positive nor negative prominence, maintaining a neutral total score of 0. The detailed scoring is available in the Appendix.

Pfizer's efforts in tackling the global COVID-19 pandemic have been met with positive media coverage, especially regarding its role in providing the vaccines (e.g., LaFraniere, 2022; Zimmer, 2021a). Prior to the pandemic, Pfizer's contributions to healthcare innovation were also a frequent topic in the news (e.g., Pollack, 2014; Thomas, 2014). However, the company's products have attracted criticism for a variety of issues, including product defects and accusations of imitating rival technologies (e.g., Jewett, 2022; Robbins and Gross, 2022). Ethical issues have also surfaced, with Pfizer being implicated in acts of fraud such as tax avoidance and breaches of anti-racketeering laws (e.g., Barro, 2014; Bloomberg News, 2011). Discussions around corporate governance have been largely neutral, focusing on changes among senior leaders, but some negative attention has been drawn due to a failed merger attempt (e.g., Hughes, 2014).

MSCI assigns Pfizer a mid-range 'A' rating within the context of 260 pharmaceutical companies (MSCI, 2023a). This evaluation by MSCI acknowledges Pfizer's need for improvement in product safety and quality (MSCI, 2023a), which corresponds with sentiment analysis findings that point to negative news regarding product defects. Despite facing ethical issues, as reported in this study and indicated by MSCI's controversy measures, Pfizer is still rated as average, suggesting that its ethical challenges are less severe than those of many competitors. Notably, Pfizer is praised as a front-runner in providing healthcare access (MSCI, 2023a), highlighted by widespread news coverage of its COVID-19 vaccine development. Pfizer is also commended for its performance in corporate governance, investment in its workforce, and reducing harmful emissions and waste (MSCI, 2023a).

#### 4.2.8. Johnson & Johnson

Established by the Johnson siblings—Robert, James, and Edward—in 1886, Johnson & Johnson has grown into a multinational powerhouse in both pharmaceuticals and medical device innovation (Johnson & Johnson, n.d.-a). Initially, their focus was on manufacturing sterile surgical products, such as sutures, absorbent cotton, and bandages (Johnson & Johnson, n.d.-a). In the present day, the company spans a wide range of health products, from everyday consumer goods like band-aids and baby powder to advanced pharmaceuticals for complex diseases and cutting-edge medical equipment for surgeries and orthopedics (Johnson & Johnson, 2023). During the COVID-19 pandemic, Johnson & Johnson expanded its health contributions by developing a vaccine to help curb the spread of the infectious disease. Under the leadership of Joaquin Duato as the CEO and Chairman, the company guides a dedicated manpower of over 150,000 people (Johnson & Johnson, n.d.-b, 2023). The global headquarters of Johnson & Johnson sit in New Brunswick, New Jersey, the same region where it began its journey (Johnson & Johnson, 2023).

Johnson & Johnson's ESG efforts are featured in 229 articles by The New York Times, less extensively covered than its competitor, Pfizer. Slightly different from Pfizer, 'Product Liability' emerges as the most scrutinized area, with 122 articles addressing concerns or developments related to the company's products and their impact on consumers. Following this, 'Social Opportunities' is covered in 48 articles, underscoring Johnson & Johnson's societal contributions. 'Business Ethics' is the subject of 37 articles, indicating a strong interest in the company's ethical standards. In contrast, 'Corporate Governance' features in only 15 articles, suggesting a more modest inquiry into the company's leadership. 'Human Capital' appears to be less of a focus, with only 5 articles. Environmental topics, namely 'Resource Stewardship' and 'Environmental Opportunities,' are the least represented in the coverage, with one article each. Neither Pfizer nor Johnson & Johnson has coverage related to 'Climate Change,' hinting at a possible underemphasis or lack of newsworthy events in this crucial ESG aspect for both companies. The distribution of the sentiments can be seen in Figure 13.

The analysis indicates a notable concentration of negative sentiment in 'Product Liability,' with 63 articles, highlighting scrutiny of the company's product safety and consumer impact. This is followed by 'Business Ethics,' where 25 articles carry a negative sentiment, suggesting a critical examination of the company's ethical conduct. Compared to Pfizer, Johnson & Johnson faces a more critical view in the category of 'Product Liability' and 'Business Ethics'. 'Social Opportunities' and 'Corporate Governance' exhibit more balanced coverage, with the majority of articles being marked as neutral. 'Human Capital' is also mostly discussed neutrally, while the environmental category, 'Resource Stewardship,' is viewed negatively in a single article, while 'Environmental Opportunities' gains a positive mention.

In assessing Johnson & Johnson's sentiment scores, it becomes clear that the company faces significant challenges in 'Product Liability,' where it has a notably negative total (normalized) score of -49 (-0.40). This is likely influenced by the company's role in the opioid epidemic, consumer litigation over product safety, and a number of product recalls (Hoffman, 2019; Hsu, 2021; Jiménez, 2021). 'Business Ethics' is another area with substantial negative sentiment, scoring -24 and normalized to -0.65, due to several lawsuits alleging improper marketing tactics, collusion, and bribery (Harris, 2011; Kanter & Thomas, 2013; Thomas, 2013). Conversely, 'Social Opportunities' shows a positive normalized (total) of 0.23, likely owing to Johnson & Johnson's efforts in developing and distributing COVID-19 vaccines and the development of new drugs (e.g., Pollack, 2012; Zimmer, 2021b). 'Corporate Governance' also carries a slight negative sentiment with a normalized (total) -0.13 (-2), slightly worse than its competitor, Pfizer. The detailed scoring is available in the Appendix.

It is important to note that the domains of 'Resource Stewardship,' 'Environmental Opportunities,' and 'Human Capital' are represented by a limited number of articles, resulting in total scores of -1, 1, and -1, with corresponding normalized scores -1.00, 1.00, and -0.20. The scant coverage in these areas introduces a degree of uncertainty to these scores, which may not fully capture the company's performance in these aspects. For instance, 'Environmental Opportunities' and 'Resource Stewardship' show very strong results, each with the highest possible normalized score. Yet, the trustworthiness of these scores is doubtful because they are based on a small amount of sentiment data.

MSCI rates Johnson & Johnson with an 'A' (Average), a ranking it shares with Pfizer within the pharmaceutical sector (MSCI, 2023a). Like Pfizer, Johnson & Johnson is recognized for its efforts to expand healthcare access, reduce harmful emissions and waste, and invest in employee development (MSCI, 2023a). The results from the sentiment score support this view, particularly highlighting Johnson & Johnson's role in distributing vaccines and developing new medications. However, the company's activities in employee development and waste management have not received much attention in the media. Regarding 'Corporate Governance,' Johnson & Johnson is considered average, while Pfizer is seen as performing better (MSCI, 2023a). Both MSCI's assessment and the sentiment scores align in identifying areas of concern for Johnson & Johnson, particularly in corporate behavior and product safety & quality, where MSCI marks it as underperforming (MSCI, 2023a). Ethical practices are a challenging area for Johnson & Johnson, with negative reports involving bribery and collusion. These ethical issues are confirmed by MSCI's controversy indicators, which note Johnson & Johnson's involvement in serious cases of bribery and fraud (MSCI, 2023a). In terms of product safety and quality, the company faces issues with product defects that cause health problems to the consumers, which are also underscored by MSCI's controversy indicators, showing the company's entanglement in significant controversies in this matter (MSCI, 2023a).

## 5. Discussions & Limitations

### 5.1. Discussions

Each industry discussed in Section 4 has its own characteristics. The technology sector, with its massive market capitalization, receives the biggest spotlight from the media. The number of articles featuring the gigantic technology companies' ESG activities is enormous compared to the coverage of other companies in other sectors. Following this, the pharmaceutical industry emerges as the second most covered sector, while the beverage sector receives the least media attention. In the technology sector, 'Business Ethics' and 'Product Liability' are the most discussed topics by the media. The ethical operations of these technology companies are mostly related to anticompetitive behavior. For instance, Apple had a case against Epic Games on how Apple abuses its power by charging developers unfairly high commissions to have a spot in Apple's App Store. Microsoft—the worst in terms of ethical conduct based on the findings, was under heavy scrutiny for its intent to acquire Activision since the deal might give an

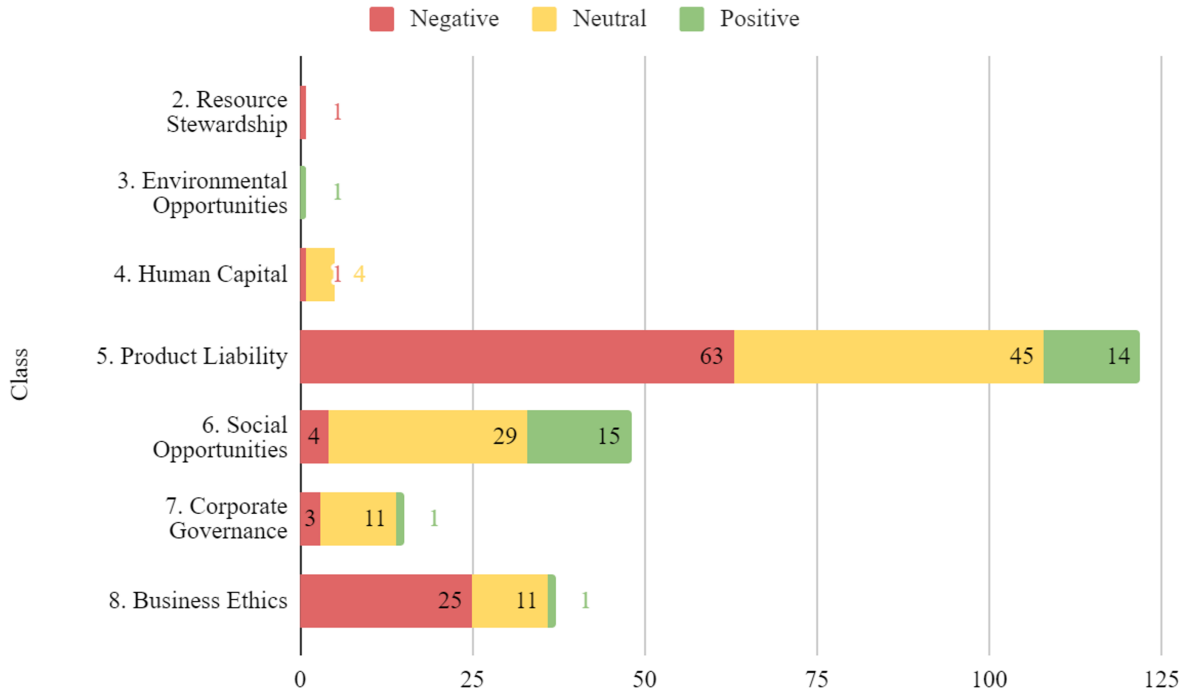


**Figure 13:** Johnson & Johnson's Sentiment Distribution Across Different ESG Categories

unfair competitive advantage to Microsoft (Weise & McCabe, 2022).

The ethical concerns for technology companies are distinct from those experienced in other industries. For other industries, a lot of issues are related to the company's unethical marketing and advertising tactics. In the beverage companies, for instance, the two companies have faced accusations of deceptive marketing strategies. The Coca-Cola Company encountered legal action over questionable claims regarding its juice products (Liptak, 2014), while PepsiCo and its subsidiary have been accused of engaging in misleading advertisements targeted at adolescents (Vega, 2011). The pharmaceutical sector also grapples with its share of ethical issues related to marketing. Pfizer was required to pay damages after being found in violation of federal racketeering laws through deceptive advertisements for one of its drugs (Bloomberg News, 2011). Johnson & Johnson faced penalties for unethical practices, including allegations of bribing doctors, which emerged during an investigation into its marketing tactics (Harris, 2011). These instances illustrate the varying ethical problems across different industries, shaped by the unique nature of the industry. Technology companies, often characterized by their large market share and dominant positions, face scrutiny over practices that could potentially harm competition, while other industries often confront ethical challenges centered around fraudulent marketing tactics.

The discourse surrounding 'Product Liability' in the technology sector also differs from that in other industries. This distinction primarily arises from the nature of the products and services offered by these industries. In the technology sector, the primary concerns highlighted in media discussions revolve around issues of data privacy and security, a direct consequence of the digital nature of the products. Taking Google as an example, it faced accusations of improperly collecting facial and voice recognition data without obtaining explicit consent from users (Hill & McCabe, 2022). Similarly, Meta was subjected to substantial fines for violating data privacy laws in Europe (Satariano, 2022b). In contrast, the pharmaceutical and beverage industries deal with more tangible products, leading to different kinds of product liability concerns. For example, in the pharmaceutical sector, a notable case is that of Johnson & Johnson, which had to recall sunscreen products due to contamination with harmful ingredients, raising serious questions about product safety and quality control (Jiménez, 2021). Meanwhile, beverage companies have been criticized for the potential health risks associated with their products (e.g., Bittman, 2012; Board, 2015).

While the technology industry has garnered media attention for ethical conduct and product liability issues, pharmaceutical companies have received prominent positive highlights in their contributions to social opportunities. Pfizer and Johnson & Johnson, not only because of their contributions in the recent pandemic, have been extensively covered by the media because of their breakthroughs in improving public health. A case in point was when Pfizer provided alternative contraceptives for developing countries beyond traditional daily birth control pills (Thomas, 2014). For Johnson & Johnson, the media acknowledged its effort in developing a new drug for prostate cancer (Pollack, 2012). In contrast, the media coverage of the beverage industry has been less favorable in terms of social contributions. Reports have high-

lighted a certain degree of hypocrisy, with companies like Coca-Cola and PepsiCo donating millions to health groups while simultaneously spending large sums to oppose public health legislation (O'Connor, 2016). In the technology sector, Microsoft is the leader in this area, with a normalized score of 0.24. However, the news articles do not necessarily report the social opportunities of its products. The articles can also include the altruistic actions of the company. For example, Microsoft pledged hundreds of millions of dollars to alleviate the housing crisis in the Seattle area (Weise, 2019). These varied narratives across industries reflect the diverse ways in which companies can impact society and how the media portrays these impacts.

Another interesting pattern that can be observed from the sentiment analysis is the limited coverage of the environmental efforts of the companies by the media. In the technology sector, which generally receives considerable media attention, a surprisingly small fraction, only about 1-3%, of the coverage is dedicated to the environmental theme. The situation appears even more pronounced in the pharmaceutical industry, where the share of environmental topics is less than 1%. Only the beverage sector receives a higher share of media coverage concerning environmental topics, by exceeding 10% of total coverage. That said, it is important to note that the news articles discussing the environmental aspect of PepsiCo predominantly feature outdated information, with seven out of eight articles being over 12 years old. In contrast, Coca-Cola has been the subject of a number of more recent news pieces. A notable example includes criticism from climate activists over Coca-Cola's sponsorship of a climate summit in Egypt in 2022, at a time when the company's production of plastics was reportedly on the rise (Engelbrecht, 2022). This pattern in media coverage might suggest that the nature of an industry significantly influences the extent of environmental reporting. The beverage industry's direct impact on resource utilization and waste production seemingly draws more media scrutiny compared to sectors like technology and pharmaceuticals.

It is crucial to underscore the importance of increased media coverage in the field of corporate environmental efforts. The current landscape, as indicated by the topic analysis results, shows a considerably low number of news articles covering companies' environmental actions. News outlets need to report more actions from the companies, both for the negative impact caused by the company and acknowledgment of positive milestones achieved by the companies. Negative coverage is essential not only for holding corporations accountable but also for informing and educating the public about the environmental impacts of various corporate actions. The media can report issues like carbon emissions, waste management, and resource utilization of the companies. Highlighting positive impacts is as crucial. This includes innovations in sustainable practices, successful implementation of environmentally friendly initiatives, significant reductions in carbon footprints, and investments in green and clean technology. Reporting on these areas can drive more informed consumer choices and potentially influence corporate policies.

Nevertheless, there is a prominent pattern in the 'Corporate Governance' theme that stands out as a resemblance across all industries. In this domain, companies from various sectors consistently exhibit scores that are either neutral or, at most, only slightly negative. Notably, Meta and Johnson & Johnson are at the lower spectrum with a score of -0.13, while Coca-Cola and PepsiCo maintain the highest scores with a neutral 0. This trend suggests a consistent but cautious media approach to corporate governance issues across various industries. A key aspect of the media coverage under this theme is the emphasis on significant changes in company leadership. For example, PepsiCo's CEO Indra Nooyi stepping down after 12 years (Creswell, 2018) and Sundar Pichai assuming the role of Alphabet's CEO are cases in point (Nicas & Wakabayashi, 2019). Such coverage is prevalent across various companies, lending a generally neutral tone to the discourse on corporate governance. Beyond leadership dynamics, the theme of corporate governance in media coverage extends to include topics like merger activities and executive pay. Instances such as Microsoft's acquisition of LinkedIn (Wingfield, 2016b) and Pfizer's proposed US$150 billion deal to buy Allergan (Merced, 2015) are illustrative of media focus on merger and acquisition cases. In another vein, executive compensation is a topic that caught media attention, particularly in the case of Coca-Cola, which faced shareholder criticism over its executive compensation (Eavis, 2014).

Compared to other research in this field, this study contributes to several key areas of improvement. J. Lee and Kim (2023) develop an ESG text classification model that allows researchers to extract ESG information from multiple sources, including reports and news articles. This study takes a leap forward by introducing a complete machine learning pipeline. It not only classifies ESG-related content but also extends to evaluating and scoring the companies' performance. In comparison to H. Kang and Kim (2022), who also utilized sentiment analysis but primarily focused on company sustainability reports, this research broadens the scope. It leverages machine learning to assess ESG performance from a third-party perspective, using news articles as the primary source. This approach provides a more balanced and external viewpoint on companies' ESG efforts compared to the self-reported nature of sustainability reports. Furthermore, this study surpasses the capabilities of the ESG-Miner tool presented by Fischbach et al. (2022). While ESG-Miner effectively identifies ESG relevance in news articles, it falls short in differentiating between the three main ESG categories. The model introduced in this study addresses this gap by providing a more detailed classification into nine distinct ESG categories, offering a finer and more insightful analysis of companies' ESG-related activities as portrayed in the media.

Furthermore, this study contributes a manually labeled ESG dataset comprising nine distinct classes, encompassing 4,500 news items. This dataset lays the groundwork for enhancing future ESG classification models. Additionally, the study illuminates the capabilities of GPT models, discovering

that for classification tasks, fine-tuning a less costly model can yield better performance than using zero-shot prompting with more expensive models. The machine learning pipeline proposed in this research effectively demonstrates how companies are represented in media reports regarding their ESG performance, highlighting the value of machine learning as a tool for assessing corporate performance. The findings offer an in-depth analysis of companies' performance across eight different ESG categories, enabling the public to measure the achievement of these companies in these areas based on media reports. This approach not only provides a clearer picture of corporate sustainability performance but also underscores the potential of machine learning in extracting meaningful insights from complex datasets.

### 5.2. Limitations

As observed in Section 4, the prevailing sentiment expressed in news articles about companies' ESG performance is predominantly neutral. This trend could be attributed to a couple of reasons. Firstly, the media tends to maintain a neutral perspective when reporting news, which is reflected in the sentiment of the articles. Secondly, the sentiment analysis model employed in this study might have an inherent bias toward classifying sentiments as neutral. The news regarding PepsiCo's commitment to reducing sugary drink calories serves as a pertinent example (Strom, 2014b). This proactive step by PepsiCo is surprisingly deemed neutral by the model, whereas it could actually indicate a positive effort from the company. This bias could be due to the nature of the training data or the algorithmic design of the model. It is important to note that the model was primarily trained on financial communication texts rather than texts specifically focused on ESG topics. This difference in the nature of the training material could significantly influence the model's tendency to categorize ESG-related sentiments as neutral. This limitation suggests that the scores derived from the sentiment analysis may not comprehensively or accurately represent the actual ESG performance of the companies.

Another significant limitation in the current research methodology arises from the lack of integration between NER and sentiment analysis. While NER effectively identifies the presence of company names within the text, the sentiment analysis component captures the overall tone. However, this approach falls short of accurately discerning the nuanced sentiment directed specifically at a company's actions within the text. The complexity of this issue becomes evident when considering mixed messages. For instance, consider an article with the title and headline of 'Beverage Companies Embrace Recycling, Until It Costs Them. Recycling is struggling in much of the United States, and companies like Coca-Cola say they are committed to fixing it.' (Corkery, 2019). The model categorizes this article as 'Negative,' aligning with the overall tone that highlights the struggles in recycling efforts. However, this classification does not do justice to the specific sentiment towards Coca-Cola's roles. The company's commitment to addressing recycling issues, a potentially positive aspect, is overlooked in this analysis. This example demonstrates how the model's inability to parse and evaluate the sentiment related to specific actions or statements of a company can lead to an incomplete understanding of the sentiment in the text.

Employing entity-level sentiment analysis for evaluating a company's performance can enhance the accu-racy of a company's sentiment. Rather than deriving a generalized sentiment from the overall tone of the text, entity-level sentiment analysis focuses on the sentiment directly associated with a particular company mentioned in the text (Sinha et al., 2023). This method is proficient at unraveling the complex layers of sentiment related to multiple entities within a single piece of text. Future research could greatly benefit from adopting this method, as demonstrated in recent studies like those of Rønningstad et al. (2023), Sinha et al. (2023), and Tang et al. (2023). Sinha et al. (2023) and Tang et al. (2023) propose innovative frameworks for extracting sentiments specifically relevant to each entity in the financial domain. Implementing entity-level sentiment analysis would allow for a more targeted and accurate assessment of how companies are perceived in regard to their ESG efforts. Nevertheless, the limitations highlighted earlier in this study underscore the necessity of integrating human evaluation into the analysis to enhance its reliability. While entity-level sentiment analysis is advancing, especially in the context of ESG-related issues, it remains a developing field. The combination of automatic and human evaluation approaches would complement each other, potentially leading to a more accurate representation of companies' actual ESG performance.

The current ESG classification model is designed to categorize text into a single ESG class. This approach presents another limitation, where it fails to account for the complexity and multifaceted nature of many news articles. Multiple ESG-related topics might be intertwined within a single piece of text. Given this limitation, a promising avenue for future research would be to explore models capable of multi-class classification. This would allow for a more nuanced and detailed understanding of the ESG topics covered in the media. Further, augmenting this multi-class approach with an entity-level sentiment analysis could offer an ever richer and more accurate analysis. By not only identifying multiple ESG themes but also associating these themes with specific entities and their corresponding sentiments. This advancement could significantly enhance the depth of ESG performance assessments derived from media analysis.

As highlighted in the previous sections, one of the limitations identified in this study is the uneven coverage of certain topics and industries, leading to an incomplete assessment of companies' ESG performance. Environmental topics, for example, there are relatively few articles discussing the efforts of companies on this front. The limited media focus contrasts with evaluation from entities like MSCI, which provide comprehensive ratings of companies' performance in environmental initiatives. Many companies, including Microsoft, Alphabet, Coca-Cola, and PepsiCo, are recognized as leaders in each group in efforts to curb carbon emissions by MSCI. This leadership, however, is not fully reflected in the senti-

ment analysis due to the sparse coverage these initiatives receive in the media. Furthermore, the extent of media coverage varies significantly among companies. For instance, beverage companies like Coca-Cola and PepsiCo receive less media attention compared to larger technology sector companies. This imbalance presents a challenge in accurately gauging the ESG performance of less-covered companies, which may be actively making significant efforts in various ESG areas. Consequently, there exists a notable discrepancy between the ESG performance ratings provided by organizations like MSCI and the findings from sentiment analysis—especially for low-covered topics and companies, a disparity that can largely attributed to the limitations in media coverage. To address this challenge, future research could expand the coverage of news articles by incorporating additional news outlets as sources of input.

Nonetheless, this study demonstrates that the machine learning pipeline can be a valuable tool for analysts and the public to gauge companies' ESG performance as portrayed in the media. While the pipeline has shown potential, it requires further refinement for more accurate assessments, as outlined earlier. Despite its current limitations, this model serves as an excellent starting point for sorting through the vast quantity of news articles. This ability to filter and analyze large datasets is a significant advantage in understanding the often-complex narrative of corporate ESG performance from the media. It is also important to acknowledge the variability in ESG ratings across different organizations is a commonly observed phenomenon as described by Berg et al. (2022) and Chatterji et al. (2016). In this study, adding a news analysis can offer additional insights. By evaluating how companies are portrayed in the media, this study provides a different angle, potentially enriching the understanding of a company's ESG efforts.

## 6. Conclusion

Driven by the significant impact of media on public perception and the growing interest in OpenAI's GPT models, I embark on an analysis of news reports to assess companies' ESG performance. This assessment is conducted using a machine learning pipeline that incorporates various models to perform a range of tasks, including NER, classification, and sentiment analysis. While this paper utilizes available open-source models for performing NER and sentiment analysis tasks, this study also examines various GPT models for classification tasks. Interestingly, the findings reveal that a fine-tuned model, even one that is more cost-effective and smaller in scale (i.e., GPT-3, ada), can outperform the zero-shot prompting capabilities of a larger, more expensive model (i.e., GPT-3.5-turbo) in classifying nine distinct ESG topics.

In evaluating the sustainability performance of corporations, this paper focuses on articles published by the New York Times between 2003 and 2022. The analysis comprises eight public corporations spanning three distinct sectors. Despite covering a broad time span, the volume of media coverage varies significantly across these sectors. The technology sector receives the most attention, while the beverage industry garners the least number of reports in the media. Based on these media reports, the sentiment of the articles related to each ESG topic is assessed. These sentiment analysis results are then utilized to generate scores for the companies, providing a measure of their sustainability performance as portrayed in the media.

The results of this study exhibit unique characteristics in media reporting for each sector. The technology sector, in particular, faces considerable media scrutiny concerning its ethical operations and product liability. Key issues here include antitrust litigations and concerns related to privacy and data security. In contrast, the pharmaceutical industry attracts more media attention regarding its contributions to public health, especially due to the COVID-19 pandemic. Meanwhile, beverage companies receive a notably higher proportion of media coverage on environmental topics than other sectors. This interest primarily focuses on the natural resource usage and waste management practices of these companies. However, the media narrative appears more uniform across the companies studied when it comes to corporate governance, with a common emphasis placed on changes in company leadership. The sentiment in the corporate governance theme tends to be balanced and generally neutral across all eight companies.

This study highlights several limitations in the current research methodology, including limited media coverage for certain topics and sectors, inaccuracies in sentiment analysis classification, the lack of integration among machine learning models, and the inability to categorize complex texts into multiple classes. Addressing these limitations could greatly enhance further research in measuring sustainability performance using machine learning algorithms. For instance, incorporating a broader range of news sources could expand media coverage, potentially addressing the issue of limited coverage observed in this study. An improvement in the current machine learning pipeline could involve the integration of various models. Combining entity-level sentiment analysis with ESG topic analysis would lead to a more accurate assessment of companies' actions.

The potential impact of this study is a significant contribution to providing a comprehensive analysis of companies' ESG performance across various industries, using news articles as the primary data source. It provides an in-depth look at how various industries and companies are portrayed in the media with respect to their ESG initiatives. This approach is particularly insightful as it reveals the distinct nature of ESG reporting across different sectors. Moreover, this research sheds light on areas where the media could potentially enhance its role in shaping public perception of ESG efforts. It suggests that a more balanced and comprehensive coverage of ESG topics, including both achievements and areas of concern, could provide a more accurate picture of companies' ESG performance. Additionally, the study implies that increased coverage of underrepresented sectors and topics could contribute to a well-informed public discourse.

## References


Allen, K. (2018). Lies, Damned Lies and ESG Rating Methodologies. *Financial Times*. https://www.ft.com/content/2e49171b-a018-3c3b-b66b-81fd7a170ab5

Alphabet Inc. (2015). Form 8-K 2015. https://www.sec.gov/Archives/edgar/data/1652044/000119312515336577/d82837d8k12b.htm

Alphabet Inc. (2019). Alphabet Management Change. https://abc.xyz/investor/news/2019/1203/

Alphabet Inc. (2022). Form 10-K 2022. https://www.sec.gov/Archives/edgar/data/1652044/000165204423000016/goog-20221231.htm

Apple Inc. (n.d.). Tim Cook. Retrieved October 31, 2023, from https://www.apple.com/leadership/tim-cook/

Apple Inc. (2022). Form 10-K 2022. https://s2.q4cdn.com/470004039/files/doc_financials/2022/q4/_10-K-2022-(As-Filed).pdf

Aue, T., Jatowt, A., & Färber, M. (2022). Predicting Companies' ESG Ratings from News Articles Using Multivariate Timeseries Analysis. *arXiv preprint arXiv:2203.00000*.

Austen, I. (2019). Trash-Picking Robots? Park Bench Monitors? Toronto Debates Tech Giant's Waterfront Plans. *The New York Times*. https://www.nytimes.com/2019/06/24/world/canada/toronto-google-sidewalk-labs.html

Barka, Z., Hamza, T., & Mrad, S. (2023). Corporate ESG Scores and Equity Market Misvaluation: Toward Ethical Investor Behavior. *Economic Modelling*, *127*, 106467. https://doi.org/10.1016/j.econmod.2023.106467

Barro, J. (2014). Pfizer's Move Poses Challenge. Here's a Solution. *The New York Times*. https://www.nytimes.com/2014/04/30/upshot/radical-solution-to-challenge-of-corporate-taxes.html

Berg, F., Fabisik, K., & Sautner, Z. (2020). Rewriting History II: The (Un)Predictable Past of ESG Ratings. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3722087

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings. *Review of Finance*, *26*(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

Bingler, J. A., Kraus, M., Leippold, M., & Webersinke, N. (2022). Cheap Talk and Cherry-Picking: What ClimateBert has to Say on Corporate Climate Risk Disclosures. *Finance Research Letters*, *47*, 102776. https://doi.org/10.1016/j.frl.2022.102776

Bittman, M. (2012). Unseasonably Warm Winter Links. *The New York Times*. https://archive.nytimes.com/bittman.blogs.nytimes.com/2012/12/05/unseasonably-warm-winter-links/

Bloomberg News. (2011). Pfizer Told to Pay $142.1 Million Over Marketing of Epilepsy Drug. *The New York Times*. https://www.nytimes.com/2011/01/29/business/29pfizer.html

Board, T. E. (2015). Opinion: Coke Tries to Sugarcoat the Truth on Calories. *The New York Times*. https://www.nytimes.com/2015/08/14/opinion/coke-tries-to-sugarcoat-the-truth-on-calories.html

Borms, S., Boudt, K., Holle, F. V., & Willems, J. (2021). Semi-supervised Text Mining for Monitoring the News About the ESG Performance of Companies. *Data Science for Economics and Finance*, 217–239. https://doi.org/10.1007/978-3-030-66891-4_10

Brackley, A., Brock, E. K., & Nelson, J. (2022, October). Rating the Raters Yet Again: Six Challenges for ESG Ratings. *ERM Sustainability Institute*. https://www.sustainability.com/thinking/rating-the-raters-yet-again-six-challenges-for-esg-ratings/

Britannica. (2023a). PepsiCo, Inc. https://www.britannica.com/topic/PepsiCo-Inc

Britannica. (2023b). The Coca-Cola Company. https://www.britannica.com/topic/The-Coca-Cola-Company

Brown, N., & Deegan, C. (1998). The Public Disclosure of Environmental Performance Information–A Dual Test of Media Agenda Setting Theory and Legitimacy Theory. *Accounting and Business Research*, *29*(1), 21–41. https://doi.org/10.1080/00014788.1998.9729564

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., & Amodei, D. (2020). Language Models are Few-Shot Learners.

Chalkidis, I., Androutsopoulos, I., & Aletras, N. (2019). Neural Legal Judgment Prediction in English.

Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (2016). Do Ratings of Firms Converge? Implications for Managers, Investors and Strategy Researchers. *Strategic Management Journal*, *37*(8), 1597–1614. https://doi.org/10.1002/smj.2407

Chen, Q., & Liu, X.-Y. (2020). Quantifying ESG Alpha Using Scholar Big Data. *Proceedings of the First ACM International Conference on AI in Finance*, 1–8. https://doi.org/10.1145/3383455.3422529

Conger, K., & Frenkel, S. (2021, March). Thousands of Microsoft Customers May Have Been Victims of Hack Tied to China.

Corkery, M. (2019). Beverage Companies Embrace Recycling, Until It Costs Them. *The New York Times*. https://www.nytimes.com/2019/07/04/business/plastic-recycling-bottle-bills.html

Creswell, J. (2018). Indra Nooyi, PepsiCo C.E.O. Who Pushed for Healthier Products, to Step Down. *The New York Times*. https://www.nytimes.com/2018/08/06/business/indra-nooyi-pepsi.html

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.

Douglas, M. R. (2023). Large Language Models.

Eavis, P. (2014). Coca-Cola, Yielding to Criticism, Revises Its Proposal for Executive Pay. *The New York Times*. https://www.nytimes.com/2014/10/02/business/coca-cola-yielding-to-criticism-revises-its-plan-for-executive-pay.html

Elias, J. (2023). Google is Asking Employees to Test Potential ChatGPT Competitors, Including a Chatbot Called 'Apprentice Bard'. *CNBC*. https://www.cnbc.com/2023/01/31/google-testing-chatgpt-like-chatbot-apprentice-bard-with-employees.html

Engelbrecht, C. (2022). Coke Is a Sponsor of the Climate Summit in Egypt. Some Activists Aren't Happy. *The New York Times*. https://www.nytimes.com/2022/11/07/climate/coca-cola-sponsor-cop27-climate-egypt.html

Fischbach, J., Adam, M., Dzhagatspanyan, V., Mendez, D., Frattini, J., Kosenkov, O., & Elahidoost, P. (2022). Automatic ESG Assessment of Companies by Mining and Evaluating Media Coverage Data: NLP Approach and Tool.

Frenkel, S. (2018). Microsoft Employees Protest Work With ICE, as Tech Industry Mobilizes Over Immigration. *The New York Times*. https://www.nytimes.com/2018/06/19/technology/tech-companies-immigration-border.html

Frenkel, S. (2022a). Meta Will Freeze Most Hiring, Zuckerberg Tells Employees. *The New York Times*. https://www.nytimes.com/2022/09/29/technology/meta-hiring-freeze.html

Frenkel, S. (2022b). Meta Is Said to Plan Significant Job Cuts This Week. *The New York Times*. https://www.nytimes.com/2022/11/06/technology/meta-layoffs.html

Gordon, C. (2023). ChatGPT Is The Fastest Growing App In The History Of Web Applications. *Forbes*. https://www.forbes.com/sites/cindygordon/2023/02/02/chatgpt-is-the-fastest-growing-ap-in-the-history-of-web-applications/?sh=10b5be27678c

Gough, N., & Chen, B. X. (2014). Groups Accuse Apple Supplier in China of Labor Violations. *The New York Times*. https://www.nytimes.com/2014/09/05/business/Apple-Supplier-Is-Accused-of-Labor-Violations.html

Goutte, S., Le-Hoang, V. P., Liu, F., & von Mettenheim, H.-J. (2023). ESG Investing: A Sentiment Analysis Approach. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4316108

Grady, D. (2020). A.I. Is Learning to Read Mammograms. *The New York Times*. https://www.nytimes.com/2020/01/01/health/breast-cancer-mammogram-artificial-intelligence.html

Guo, T., Jamet, N., Betrix, V., Piquet, L.-A., & Hauptmann, E. (2020). ESG2Risk: A Deep Learning Framework from ESG News to Stock Volatility Prediction.

Hall, M., & Hosch, W. L. (2023). Google. https://www.britannica.com/topic/Google-Inc

Hammami, A., & Hendijani Zadeh, M. (2019). Audit quality, media coverage, environmental, social, and governance disclosure and firm investment efficiency. *International Journal of Accounting & Information Management*, *28*(1), 45–72. https://doi.org/10.1108/IJAIM-03-2019-0041

Hardy, Q. (2016). Google Says It Will Run Entirely on Renewable Energy in 2017. *The New York Times*. https://www.nytimes.com/2016/12

/06/technology/google-says-it-will-run-entirely-on-renewable-energy-in-2017.html

Harris, G. (2011). Johnson & Johnson Settles Bribery Complaint for $70 Million in Fines. *The New York Times*. https://www.nytimes.com/2011/04/09/business/09drug.html

Hartzmark, S. M., & Sussman, A. B. (2019). Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows. *The Journal of Finance*, *74*(6), 2789–2837. https://doi.org/10.1111/jofi.12841

Hill, K., & McCabe, D. (2022). Texas Sues Google for Collecting Biometric Data Without Consent. *The New York Times*. https://www.nytimes.com/2022/10/20/technology/texas-google-privacy-lawsuit.html

Hirai, A., Brady, A., & Partners, S. (2021, July). Managing ESG Data and Rating Risk. https://corpgov.law.harvard.edu/2021/07/28/managing-esg-data-and-rating-risk/

Hoffman, J. (2019). Johnson & Johnson Ordered to Pay $572 Million in Landmark Opioid Trial. *The New York Times*. https://www.nytimes.com/2019/08/26/health/oklahoma-opioids-johnson-and-johnson.html

Holpuch, A. (2022). Two Women Sue Apple Over AirTag Stalking. *The New York Times*. https://www.nytimes.com/2022/12/06/business/apple-airtag-lawsuit.html

Honnibal, M., & Montani, I. (2017). spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing.

Hsu, T. (2021). Black Women's Group Sues Johnson & Johnson Over Talc Baby Powder. *The New York Times*. https://www.nytimes.com/2021/07/27/business/johnson-baby-powder-black-women.html

Huang, A. (2022). Description of 9-class Environmental, Social and Governance (ESG) Classification. https://www.allenhuang.org/uploads/2/6/5/5/26555246/esg_9-class_descriptions.pdf

Huang, A. H., Wang, H., & Yang, Y. (2023). FinBERT: A Large Language Model for Extracting Information from Financial Text. *Contemporary Accounting Research*, *40*(2), 806–841. https://doi.org/10.1111/1911-3846.12832

Hughes, C. (2014). The Lessons of a Drug Maker's Failed Deal. *The New York Times*. https://dealbook.nytimes.com/2014/05/27/the-lessons-of-a-failed-drug-deal/

Ilango, H. (2023, May). An Unregulated ESG Rating System Reveals Its Flaws. *Institute for Energy Economics & Financial Analysis*.

Isaac, M. (2018). Instagram's Co-Founders to Step Down From Company. *The New York Times*. https://www.nytimes.com/2018/09/24/technology/instagram-cofounders-resign.html

Isaac, M. (2021). Facebook and Its Apps Suffer Another Outage. *The New York Times*. https://www.nytimes.com/2021/10/08/technology/facebook-whatsapp-instagram-down.html

Jewett, C. (2022). Pfizer Recalls Some Blood Pressure Drugs, Citing Cancer Risk. *The New York Times*. https://www.nytimes.com/2022/03/23/health/pfizer-recall-blood-pressure-drug-cancer.html

Jiménez, J. (2021). Johnson & Johnson Recalls Sunscreen Because of Benzene Traces. *The New York Times*. https://www.nytimes.com/2021/07/14/us/johnson-johnson-sunscreen-recall-aveeno-neutrogena.html

Johnson & Johnson. (n.d.-a). Our Beginning. Retrieved October 31, 2023, from https://ourstory.jnj.com/our-beginning

Johnson & Johnson. (n.d.-b). Our Leadership Team. Retrieved October 31, 2023, from https://www.jnj.com/our-leadership-team

Johnson & Johnson. (2023). Form 10-K 2022. https://johnsonandjohnson.gcs-web.com/static-files/06bc3388-603b-4768-bf95-e6d43fda9fd3

Johnston, M. (2022). Biggest Companies in the World by Market Cap. *Investopedia*. https://www.investopedia.com/biggest-companies-in-the-world-by-market-cap-5212784

Kaddour, J., Harris, J., Mozes, M., Bradley, H., Raileanu, R., & McHardy, R. (2023). Challenges and Applications of Large Language Models.

Kang, C., Isaac, M., & Popper, N. (2019). Facebook's Zuckerberg, Accused of Lying, Withstands a Washington 'Beating'. *The New York Times*. https://www.nytimes.com/2019/10/23/technology/facebook-zuckerberg-libra-congress.html

Kang, H., & Kim, J. (2022). Analyzing and Visualizing Text Information in Corporate Sustainability Reports Using Natural Language Processing Methods. *Applied Sciences*, *12*(11), 5614. https://doi.org/10.3390/app12115614

Kanter, J., & Thomas, K. (2013). Europe Says Drug Makers Paid to Delay a Generic. *The New York Times*. https://www.nytimes.com/2013/02/01/business/global/eu-says-drug-makers-paid-to-delay-generic-version.html

Kelly, S. M. (2023). ChatGPT Passes Exams from Law and Business Schools. *CNN Business*. https://edition.cnn.com/2023/01/26/tech/chatgpt-passes-exams/index.html

Krappel, T., Bogun, A., & Borth, D. (2021). Heterogeneous Ensemble for ESG Ratings Prediction.

LaFraniere, S. (2022). New Booster Shot Targets Covid Variants More Effectively, Pfizer Says. *The New York Times*. https://www.nytimes.com/2022/11/04/us/politics/covid-booster-pfizer.html

Lattman, P. (2012). Former Coca-Cola Bottling Executive Charged With Insider Trading. *The New York Times*. https://dealbook.nytimes.com/2012/03/08/s-e-c-charges-former-coca-cola-bottling-executive-with-insider-trading/

Lee, H., Choi, J., Kwon, S., & Jung, S. (2023). EaSyGuide: ESG Issue Identification Framework Leveraging Abilities of Generative Large Language Models.

Lee, J., & Kim, M. (2023). ESG Information Extraction with Cross-Sectoral and Multi-Source Adaptation Based on Domain-Tuned Language Models. *Expert Systems with Applications*, *221*, 119726. https://doi.org/10.1016/j.eswa.2023.119726

Leite, B. J., & Uysal, V. B. (2023). Does ESG Matter to Investors? ESG Scores and the Stock Price Response to New Information. *Global Finance Journal*, *57*, 100851. https://doi.org/10.1016/j.gfj.2023.100851

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2019). BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension.

Li, H. (2022). Language Models: Past, Present, and Future. *Communications of the ACM*, *65*(7), 56–63. https://doi.org/10.1145/3490443

Li, Y., Li, Z., Zhang, K., Dan, R., Jiang, S., & Zhang, Y. (2023). ChatDoctor: A Medical Chat Model Fine-Tuned on a Large Language Model Meta-AI (LLaMA) Using Medical Domain Knowledge.

Linzmayer, O. (2004). *Apple Confidential 2.0: The Definitive History of the World's Most Colorful Company*. No Starch Press.

Liptak, A. (2014). Coke Can Be Sued by Rival Over Juice Claim, Court Says. *The New York Times*. https://www.nytimes.com/2014/06/13/business/supreme-court-says-coca-cola-can-be-sued-by-Pom-Wonderful.html

Lohr, S. (2020). Slack Accuses Microsoft of Illegally Crushing Competition. *The New York Times*. https://www.nytimes.com/2020/07/22/technology/slack-microsoft-antitrust.html

Luccioni, A., Baylor, E., & Duchene, N. (2020). Analyzing Sustainability Reports Using Natural Language Processing.

Mac, R. (2022). Lawsuit says Meta shares blame in the killing of a federal guard. *The New York Times*. https://www.nytimes.com/2022/01/06/technology/meta-facebook-lawsuit-security-guard.html

Malo, P., Sinha, A., Takala, P., Korhonen, P., & Wallenius, J. (2013). Good Debt or Bad Debt: Detecting Semantic Orientations in Economic Texts.

Markoff, J. (2003). TECHNOLOGY; RealNetworks Accuses Microsoft Of Restricting Competition. *The New York Times*. https://www.nytimes.com/2003/12/19/business/technology-realnetworks-accuses-microsoft-of-restricting-competition.html

McCombs, M., & Reynolds, A. (2002). *Media Effects* (J. Bryant, D. Zillmann, J. Bryant, & M. B. Oliver, Eds.). Routledge. https://doi.org/10.4324/9781410602428

Mehra, S., Louka, R., & Zhang, Y. (2022). ESGBERT: Language Model to Help with Classification Tasks Related to Companies' Environmental, Social, and Governance Practices. *Embedded Systems and Applications*, 183–190. https://doi.org/10.5121/csit.2022.120616

Mele, C. (2016). Bags of Cocaine Worth $56 Million Are Found at Coca-Cola Factory in France. *The New York Times*. https://www.nytimes.co

m/2016/09/02/world/europe/bags-of-cocaine-worth-56-million-are-found-at-coca-cola-factory-in-france.html
Merced, M. J. d. l. (2015). Pfizer and Allergan Said to Be Near Merger for Up to $150 Billion. *The New York Times*. https://www.nytimes.com/2015/11/19/business/dealbook/pfizer-allergan-deal-for-up-to-150-billion-is-said-to-be-close-to-complete.html
Meta Platforms Inc. (n.d.). Mark Zuckerberg, Founder, Chairman and Chief Executive Officer. Retrieved October 31, 2023, from https://about.meta.com/media-gallery/executives/mark-zuckerberg/
Meta Platforms Inc. (2022). Form 10-K 2022. https://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/e574646c-c642-42d9-9229-3892b13aabfb.pdf
Microsoft. (n.d.-a). Facts About Microsoft. Retrieved February 12, 2023, from https://news.microsoft.com/facts-about-microsoft/
Microsoft. (n.d.-b). Satya Nadella. Retrieved October 31, 2023, from https://news.microsoft.com/exec/satya-nadella/
Microsoft. (2023). Form 10-K 2022. https://microsoft.gcs-web.com/node/31736/html
Morales, C. (2022). Restaurants Face an Extortion Threat: A Bad Rating on Google. *The New York Times*. https://www.nytimes.com/2022/07/11/dining/google-one-star-review-scam-restaurants.html
MSCI. (2023a). ESG Ratings & Climate Search Tool. https://www.msci.com/our-solutions/esg-investing/esg-ratings-climate-search-tool
MSCI. (2023b). ESG Ratings Methodology.
Newman, A. A. (2011). Good/Corps Aims to Help Business Meet Social Goals. *The New York Times*. https://www.nytimes.com/2011/05/13/business/media/13adco.html
Nguyen, Q., Diaz-Rainey, I., & Kuruppuarachchi, D. (2020). Predicting Corporate Carbon Footprints for Climate Finance Risk Analyses: A Machine Learning Approach. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3617175
Nicas, J. (2019). Apple Removes App That Helps Hong Kong Protesters Track the Police. *The New York Times*. https://www.nytimes.com/2019/10/09/technology/apple-hong-kong-app.html
Nicas, J., Browning, K., & Griffith, E. (2020). Fortnite Creator Sues Apple and Google After Ban From App Stores. *The New York Times*. https://www.nytimes.com/2020/08/13/technology/apple-fortnite-ban.html
Nicas, J., & Wakabayashi, D. (2019). Era Ends for Google as Founders Step Aside From a Pillar of Tech. *The New York Times*. https://www.nytimes.com/2019/12/03/technology/google-alphabet-ceo-larry-page-sundar-pichai.html
Nolen, J. L. (2023). Pfizer, Inc. https://www.britannica.com/topic/Pfizer-Inc
Nugent, T., Stelea, N., & Leidner, J. L. (2020). Detecting ESG topics using domain-specific language models and data augmentation approaches.
O'Connor, A. (2016). Coke and Pepsi Give Millions to Public Health, Then Lobby Against It. *The New York Times*. https://www.nytimes.com/2016/10/10/well/eat/coke-and-pepsi-give-millions-to-public-health-then-lobby-against-it.html
OpenAI. (n.d.). Fine-tuning (Legacy). Retrieved October 27, 2023, from https://platform.openai.com/docs/guides/legacy-fine-tuning/fine-tuning
OpenAI. (2022, October). Introducing ChatGPT. https://openai.com/blog/chatgpt
Papers With Code. (n.d.). Text Classification on financial_phrasebank. Retrieved October 18, 2023, from https://paperswithcode.com/sota/text-classification-on-financial-phrasebank
PepsiCo Inc. (n.d.). Our Leadership. Retrieved October 31, 2023, from https://www.pepsico.com/who-we-are/leadership
PepsiCo Inc. (2022). Form 10-K 2022. https://investors.pepsico.com/docs/default-source/investors/q4-2022/q4-2022-form-10k_hmielz4d40rd4s16.pdf
Peric, L., Mijic, S., Stammbach, D., & Ash, E. (2020). Legal Language Modeling with Transformers. *CEUR Workshop Proceedings*. https://doi.org/10.3929/ethz-b-000456079
Pfizer Inc. (n.d.). Executive Leadership. Retrieved October 31, 2023, from https://www.pfizer.com/about/people/executives
Pfizer Inc. (2022). Form 10-K 2022. https://www.sec.gov/ix?doc=/Archives/edgar/data/78003/000007800323000024/pfe-20221231.htm
Picker, L. (2016). Billionaire Investor Nelson Peltz Sells Stake in PepsiCo. *The New York Times*. https://www.nytimes.com/2016/05/14/business/dealbook/billionaire-investor-nelson-peltz-sells-stake-in-pepsico.html
Polignano, M., Bellantuono, N., Lagrasta, F. P., Caputo, S., Pontrandolfo, P., & Semeraro, G. (2022). An NLP Approach for the Analysis of Global Reporting Initiative Indexes from Corporate Sustainability Reports. *Proceedings of the First Computing Social Responsibility Workshop within the 13th Language Resources and Evaluation Conference*, 1–8. https://aclanthology.org/2022.csrnlp-1.1
Pollack, A. (2012). Trial Shows Benefit in Using Prostate Cancer Drug Early. *The New York Times*. https://www.nytimes.com/2012/03/09/business/trial-shows-benefit-in-earlier-use-of-zytiga-for-prostate-cancer.html
Pollack, A. (2014). Guarded Optimism After Breast Cancer Drug Shows Promising Results. *The New York Times*. https://www.nytimes.com/2014/04/07/business/breast-cancer-drug-shows-groundbreaking-results.html
Principles for Responsible Investment (PRI). (2023). PRI Annual Report: Responsible investment ecosystems. https://www.unpri.org/annual-report-2023/responsible-investment-ecosystems#fn_link_1
PwC. (2022). Asset and wealth management revolution 2022: Exponential expectations for ESG. https://www.pwc.com/gx/en/financial-services/assets/pdf/pwc-awm-revolution-2022.pdf
Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving Language Understanding by Generative Pre-Training. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf
Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2019). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer.
Robbins, R., & Gross, J. (2022). Moderna Sues Pfizer and BioNTech Over Covid Vaccine Technology. *The New York Times*. https://www.nytimes.com/2022/08/26/business/moderna-covid-vaccine-lawsuit.html
Rønningstad, E., Velldal, E., & Øvrelid, L. (2023). Entity-Level Sentiment Analysis (ELSA): An exploratory task survey. https://arxiv.org/pdf/2304.14241.pdf
Satariano, A. (2020). Facebook Loses Antitrust Decision in Germany Over Data Collection. *The New York Times*. https://www.nytimes.com/2020/06/23/technology/facebook-antitrust-germany.html
Satariano, A. (2021). Apple's App Store Draws E.U. Antitrust Charge. *The New York Times*. https://www.nytimes.com/2021/04/30/technology/apple-antitrust-eu-app-store.html
Satariano, A. (2022a). A secret ad deal between Google and Meta is under scrutiny in Europe. *The New York Times*. https://www.nytimes.com/2022/03/11/business/google-meta-eu-britain-inquiry.html
Satariano, A. (2022b). Meta Fined $275 Million for Breaking E.U. Data Privacy Law. *The New York Times*. https://www.nytimes.com/2022/11/28/business/meta-fine-eu-privacy.html
Satariano, A., & Frenkel, S. (2022). Oversight Board Criticizes Meta for Preferential Treatment. *The New York Times*. https://www.nytimes.com/2022/12/06/technology/meta-preferential-treatment.html
Scheiber, N. (2020). Labor Board Accuses Google Contractor of Violating Union Rights. *The New York Times*. https://www.nytimes.com/2020/10/08/business/google-nlrb-hcl-union.html
Schwartz, J. (2015). Coca-Cola Says It's Close to Water Replenishment Goal. *The New York Times*. https://www.nytimes.com/2015/08/26/business/coca-cola-expects-to-reach-its-water-replenishment-goal-5-years-early.html
Scott, M. (2015). Skype Service Problems for Some Users Worldwide. *The New York Times*. https://www.nytimes.com/2015/09/22/technology/skype-service-disrupted-for-some-users-worldwide.html
Sejnowski, T. J. (2023). Large Language Models and the Reverse Turing Test. *Neural Computation*, *35*(3), 309–342. https://doi.org/10.1162/neco_a_01563

Serafeim, G., & Yoon, A. (2021). Stock Price Reactions to ESG News: The Role of ESG Ratings and Disagreement. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3765217

Singer, N. (2018). Apple, in Sign of Health Ambitions, Adds Medical Records Feature for iPhone. *The New York Times*. https://www.nytimes.com/2018/01/24/technology/apple-iphone-medical-records.html

Singhal, K., Azizi, S., Tu, T., Mahdavi, S. S., Wei, J., Chung, H., Scales, N., Tanwani, A., Cole-Lewis, H., Pfohl, S., Payne, P., Seneviratne, M., Gamble, P., Kelly, C., Scharli, N., Chowdhery, A., Mansfield, P., Arcas, B. A. Y., Webster, D., & Natarajan, V. (2022). Large Language Models Encode Clinical Knowledge.

Singhal, K., Tu, T., Gottweis, J., Sayres, R., Wulczyn, E., Hou, L., Clark, K., Pfohl, S., Cole-Lewis, H., Neal, D., Schaekermann, M., Wang, A., Amin, M. F., Lachgar, S., Mansfield, P., Prakash, S., Green, B., Dominowska, E., Arcas, B. A. Y., & Natarajan, V. (2023). Towards Expert-Level Medical Question Answering with Large Language Models.

Sinha, A., Kedas, S., Kumar, R., & Malo, P. (2023). SEntFiN 1.0: Entity-Aware Sentiment Analysis for Financial News. https://doi.org/10.1002/asi.24634

Southall, A. (2015). One Dead After Truck Hits Apartment Building in the Bronx. *The New York Times*. https://www.nytimes.com/2015/10/13/nyregion/one-dead-after-truck-hits-building-scaffolding-in-the-bronx.html

Starks, L. T. (2023). Presidential Address: Sustainable Finance and ESG Issues—Value versus Values. *The Journal of Finance*, *78*(4), 1837–1872. https://doi.org/10.1111/jofi.13255

Strom, S. (2010). Pepsi Refresh Contestant Claims Rules Were Broken. *The New York Times*. https://www.nytimes.com/2010/10/01/business/01pepsi.html

Strom, S. (2011). PepsiCo to Foster Chickpeas in Ethiopia. *The New York Times*. https://www.nytimes.com/2011/09/21/business/pepsicos-chick-pea-plan-includes-taking-on-famine.html

Strom, S. (2014a). Coca-Cola to Remove an Ingredient Questioned by Consumers. *The New York Times*. https://www.nytimes.com/2014/05/06/business/coca-cola-to-remove-an-ingredient-questioned-by-consumers.html

Strom, S. (2014b). Soda Makers Coca-Cola, PepsiCo and Dr Pepper Join in Effort to Cut Americans' Drink Calories. *The New York Times*. https://www.nytimes.com/2014/09/24/business/big-soda-companies-agree-on-effort-to-cut-americans-drink-calories.html

Strom, S. (2015). Coca-Cola to Cut Up to 1,800 Jobs. *The New York Times*. https://www.nytimes.com/2015/01/09/business/coca-cola-announces-plan-to-cut-1600-to-1800-jobs.html

Sun, X., Li, X., Li, J., Wu, F., Guo, S., Zhang, T., & Wang, G. (2023). Text Classification via Large Language Models.

Tang, Y., Yang, Y., Huang, A. H., Tam, A., & Tang, J. Z. (2023). FinEntity: Entity-level Sentiment Classification for Financial Texts. https://arxiv.org/pdf/2310.12406.pdf

The Coca-Cola Company. (n.d.-a). James Quincey. Retrieved October 31, 2023, from https://www.coca-colacompany.com/about-us/leadership/james-quincey

The Coca-Cola Company. (n.d.-b). Our Company. Retrieved October 31, 2023, from https://www.coca-colacompany.com/about-us

The Coca-Cola Company. (2022). Form 10-K 2022. https://investors.coca-colacompany.com/filings-reports/all-sec-filings/content/0000021344-23-000011/ko-20221231.htm

The GDELT Project. (n.d.-a). The GDELT Project. Retrieved October 15, 2023, from https://www.gdeltproject.org/

The GDELT Project. (n.d.-b). The GDELT Story. Retrieved October 15, 2023, from https://www.gdeltproject.org/about.html

The New York Times Company. (2023). The New York Times Company Reports Second-Quarter 2023 Results.

Thomas, K. (2013). J.&J. to Pay $2.2 Billion in Risperdal Settlement. *The New York Times*. https://www.nytimes.com/2013/11/05/business/johnson-johnson-to-settle-risperdal-improper-marketing-case.html

Thomas, K. (2014). Pfizer and Aid Groups Team Up on Contraceptive for Developing World. *The New York Times*. https://www.nytimes.com/2014/11/14/business/pfizer-and-aid-groups-team-up-on-depo-provera-for-developing-world.html

Tracy, M. (2021). A West Virginia newspaper company is suing Google and Facebook over online ads. *The New York Times*. https://www.nytimes.com/live/2021/01/29/business/us-economy-coronavirus

United Nations Global Compact. (2004). Who Cares Wins: Connecting Financial Markets to a Changing World. https://www.unepfi.org/fileadmin/events/2004/stocks/who_cares_wins_global_compact_2004.pdf

Vance, A. (2010). Microsoft and New York in Software Deal. *The New York Times*. https://www.nytimes.com/2010/10/21/technology/21soft.html

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need.

Vega, T. (2011). Complaint Accuses Pepsi of Deceptive Marketing. *The New York Times*. https://archive.nytimes.com/mediadecoder.blogs.nytimes.com/2011/10/19/complaint-accuses-pepsi-of-deceptive-marketing/

Venigalla, A., Frankle, J., & Carbin, M. (2022, December). BioMedLM: a Domain-Specific Large Language Model for Biomedical Text. *Mosaic ML*. https://www.mosaicml.com/blog/introducing-pubmed-gpt

Wakabayashi, D. (2021). Google Temps Fought Loss of Pandemic Bonus. And Won. *The New York Times*. https://www.nytimes.com/2021/11/05/technology/google-workers.html

Webersinke, N., Kraus, M., Bingler, J. A., & Leippold, M. (2021). ClimateBert: A Pretrained Language Model for Climate-Related Text.

Weise, K. (2019). Microsoft Pledges $500 Million for Affordable Housing in Seattle Area. *The New York Times*. https://www.nytimes.com/2019/01/16/technology/microsoft-affordable-housing-seattle.html

Weise, K., & McCabe, D. (2022). F.T.C. Sues to Block Microsoft's $69 Billion Acquisition of Activision. *The New York Times*. https://www.nytimes.com/2022/12/08/technology/ftc-microsoft-activision.html

Wingfield, N. (2016a). Microsoft Cutting 1,850 Jobs in Smartphone Unit. *The New York Times*.

Wingfield, N. (2016b). With LinkedIn, Microsoft Looks to Avoid Past Acquisition Busts. *The New York Times*. https://www.nytimes.com/2016/12/08/technology/with-linkedin-microsoft-looks-to-avoid-past-acquisition-busts.html

Woo, E. (2022). Fitbit recalls more than one million smart watches over a burn risk. *The New York Times*. https://www.nytimes.com/2022/03/02/business/fitbit-ionic-watch-recall.html

Wu, S., Irsoy, O., Lu, S.-w., Dabravolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., & Mann, G. (2023). BloombergGPT: A Large Language Model for Finance.

Zachary, G. P., & Hall, M. (2023, February). Microsoft Corporation. https://www.britannica.com/topic/Microsoft-Corporation

Zhuang, Y. (2022). New Crack in Apple's Armor as Dozens Strike at Its Stores in Australia. *The New York Times*. https://www.nytimes.com/2022/10/17/business/apple-store-strike-australia.html

Zimmer, C. (2021a). Pfizer Says Its Vaccine Works Against Key Mutation in Contagious Variants. *The New York Times*. https://www.nytimes.com/2021/01/08/health/pfizer-covid-vaccine-variant-mutation.html

Zimmer, C. (2021b). J&J's Booster Shot Provides Strong Protection against Severe Disease from Omicron, a Study Says. *The New York Times*. https://www.nytimes.com/2021/12/30/health/johnson-vaccine-booster-omicron.html