A thesis submitted to the Department of Environmental Sciences and Policy of Central European University in part fulfilment of the Degree of Master of Science

A Novel Natural Language Processing Framework for Identifying Corporate Greenwashing

Blanka, TOTH

June, 2024

Vienna

CEU eTD Collection

Notes on copyright and the ownership of intellectual property rights:

(1) Copyright in text of this thesis rests with the Author. Copies (by any process) either in full, or of extracts, may be made only in accordance with instructions given by the Author and lodged in the Central European University Library. Details may be obtained from the Librarian. This page must form part of any such copies made. Further copies (by any process) of copies made in accordance with such instructions may not be made without the permission (in writing) of the Author.

(2) The ownership of any intellectual property rights which may be described in this thesis is vested in the Central European University, subject to any prior agreement to the contrary, and may not be made available for use by third parties without the written permission of the University, which will prescribe the terms and conditions of any such agreement.

(3) For bibliographic and reference purposes this thesis should be referred to as:

Toth, B. 2024. *A Novel Natural Language Processing Framework for Identifying Corporate Greenwashing.* Master of Science thesis, Central European University, Vienna.

Further information on the conditions under which disclosures and exploitation may take place is available from the Head of the Department of Environmental Sciences and Policy, Central European University.

CEU eTD Collection

Authors Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Blanka TOTH

CEU eTD Collection

CENTRAL EUROPEAN UNIVERSITY

ABSTRACT OF THESIS submitted by:

Blanka TOTH

for the degree of Master of Science and entitled: A Novel Natural Language Processing Framework for Identifying Corporate Greenwashing

Month and Year of submission: June, 2024.

This thesis introduces a natural language processing (NLP) based framework employing the "Bidirectional Encoder Representations from Transformers" (BERT) deep learning model to detect greenwashing in corporate communication. This framework, albeit resting on predefined literature, is unique in its approach and methodology. The model utilizes multiple definitions of greenwashing created based on a literature review (deception and misinformation, misleading communication, selective disclosure, and greenwashing as decoupling), to capture the nuanced language of corporate sustainability communication. Validated on a diverse dataset from Google News article titles of the Dow Jones Industrial Average companies, the model was evaluated using precision, recall, F1-score, and Chi-squared test, and its results proved to be statistically significant. The model identified verified cases of greenwashing with a precision of 91.7%, which is considered high, but the lower recall rate of 55% signals that detecting all forms of greenwashing remains challenging. However, the definitions that were created based on an extensive literature review significantly enhanced the model's ability to differentiate and detect different greenwashing approaches, contributing both to theoretical understanding and practical applications. The proposed framework is an improved version of existing greenwashing detection methods that rely on keyword search as a rule of thumb, thereby supporting regulators to consumers by advocating for the more transparent disclosure of corporate environmental practices and environmental impact.

Keywords: greenwashing detection, natural language processing, corporate greenwashing, transformers, sustainability reporting

CEU eTD Collection

Acknowledgements

I would like to express my gratitude to a number of people and entities whose support was invaluable during my study. Firstly, I my deepest thanks to the Department of Environmental Sciences and Policy, who have enriched my life with diverse perspectives and experiences. I am also deeply inspired by nature, which has been a constant source of motivation and reflection for me. It was the beauty and complexity of the natural world that was driving me to explore the intersections between environmental sustainability and technology.

The research is dedicated to Marcell, who has been my cornerstone throughout the writing journey. His support, encouragement, and belief in my capabilities have been my mainstay against all challenges.

I am grateful to all these forces and my family, for guiding me through my academic journey, and I look forward to carrying these valuable lessons into my future.

CEU eTD Collection

Table of Contents

CEU eTD Collection

CEU eTD Collection

List of Figures and Tables

CEU eTD Collection

List of Abbreviations

Abbreviation	Full Term
AI	Artificial Intelligence
NLP	Natural Language Processing
BERT	Bidirectional Encoder Representations from Transformers
ESG	Environmental Social Governance
CS	Corporate Sustainability
CSR	Corporate Social Responsibility
TCFD	Task Force on Climate-Related Financial Disclosures
DIJA	Dow Jones Industrial Average
SDG	Sustainable Development Goals
UN	United Nations
EPA	Environmental Protection Agency

CEU eTD Collection

1. Chapter: Introduction

1.1 Background

In a world where being “green” has become a badge of honor for businesses, the lines between genuine environmental practices and greenwashing tactics have never been blurrier. As sustainability reports saturate the business landscape and media headlines praise eco-friendly projects increasingly, consumers are left to navigate a sea of opaque and deceptive information. This raises the question: how can we develop effective models for detecting greenwashing practices?

There is a growing emphasis on environmental responsibility in today's corporate world. Sustainability reports and the Task Force on Climate-Related Financial Disclosures (TCFD) are becoming increasingly universal, media stories highlight environmentally conscious projects, and businesses take immense pride in being "green." However, as more companies report on their environmental impact (these reports are still rather mandatory than voluntary), scrutiny is increasing on the trustworthiness of these reports and claims, raising concerns about potential greenwashing. More and more information are available, but it is difficult for consumers to navigate through the pile of data and learn about the verified, real environmental impact of companies. Distinguishing between genuine sustainability efforts and mere greenwashing tactics is a challenge, since the current landscape is information-rich, but unclear.

Although lacking a legal definition, greenwashing involves the deceptive tactics employed by companies to highlight their environmental commitments, these often involve the dissemination or masking of misleading information, or downplaying potentially harmful activities (Doyle,

CEU eTD Collection

2016). The gravity of greenwashing is because in most cases, it creates a false promise of corporate environmental responsibility, which can undermine genuine efforts towards sustainability agendas (Addisu 2021).

Companies risk their reputation and compromise the integrity of the broader sustainability movement when they make deceptive environmental claims. Therefore, greenwashing is a key concern among scholars researching climate change (Nguyen 2023) and is being researched increasingly often. Greenwashing is a multi-faceted phenomenon and is a problem in the fields of marketing and advertising, business and sustainability reporting, corporate social responsibility, climate change research, consumer behavior, and even law (Bernini and La Rosa 2023). Greenwashing is also subjective - what one person considers greenwashing another might see as legitimate marketing. There is no clear line between exaggeration and deception. Because of this, it is almost impossible to verify claims of companies related to environmental impacts. In 2020 the European Commission found that "53.3% of examined environmental claims in the EU were found to be vague, misleading or unfounded and 40% were unsubstantiated" (European Commission 2020). The lack of uniform guidelines for companies making voluntary green claims encourages greenwashing in the EU. In March of 2023, the Commission proposed the Green Claims Directive to provide consumers with more information about what really is being marketed as sustainable (European Commission 2023). After the implementation of the directive, before a company could label a product "recycled" or "CO2 compensated", the claim would be scientifically backed and verified. Another aim of the directive is to focus on and uncover those environmental impacts that are most relevant for the given company, because today there is a trend of focusing on areas that can easily be exaggerated but are in fact only a minor part of the environmental consequences of the business operation. In contrast to current EU legislation,

CEU eTD Collection

consumer organizations would have the legal capacity to defend the rights of consumers in case of a suspected misleading claim.

While the greenwashing situation is far from being eliminated in Europe, possible mitigating regulation is underway. In the United States, however, the situation is larger-scale and much more devastating. In the U.S., there is a lack of federal legislation targeting misleading environmental claims, leaving the issue primarily to the Federal Trade Commission (FTC). The FTC has Green Guides that help sellers in steering clear of misleading consumers with false environmental claims. However, greenwashing in the United States is addressed through case-by-case enforcement actions. This motivates companies to stay under the radar or have a deceptive strategy that has not been reported before. Without standardized penalties, companies are not discouraged to use greenwashing tactics, especially if potential financial gains outweigh the chances of getting caught.

In 2010, TerraChoice (a research group focused on environmental marketing), found that 95% of 5300 green product claims in the US were misleading (Tomasulo 2010). In 2023, the Wall Street Journal conducted a survey with more than 1500 executives and found that 60% of companies acknowledge overstating their sustainability data or aspirations (The Wall Street Journal 2023). Another result of this survey was that only 22% of businesses have measured the results of their sustainability initiatives, even though 94% of them claim to have begun creating or implementing some sort of sustainability plan. These numbers show the need for stricter legislation and enforcement measures in the United States to promote transparency in corporate environmental practices. What could be even more promising, and an immediate solution is to address the problem at its source and detect greenwashing early on.

CEU eTD Collection

1.2 Literature gap, research question, hypothesis

The research gap addressed in this thesis is AI greenwashing detection with a focus on greenwashing detection methods using natural language processing (NLP). While there is extensive research on greenwashing classifications and various detection methods, little research has been done on using AI for more effective greenwashing methods.

The current attempts to detect greenwashing rely heavily on keyword matching, sentiment analysis of limited text data, or using an arbitrarily selected baseline greenwashing definition (Santos 2023). Although these approaches do advance the field of environmental sciences, they most of the times result in partially misclassified data (false positives or false negatives), and struggle to adapt to the constantly changing language of greenwashing. So far, no approach or framework has resulted in what can be a golden source for greenwashing detection.

The current study proposes a framework that combines traditional search methods with advanced NLP techniques. The framework utilizes the BERT (Bidirectional Encoder Representations from Transformers) deep learning model with pre-defined greenwashing statements obtained from an extensive literature review. By incorporating these custom definitions of different greenwashing approaches, the model aims catch the nuances of greenwashing language and improve the precision and recall of detecting such practices, thereby contributing to better transparency and accountability in corporate sustainability efforts.

This research attempts to explore and address the following research question:

"Can advanced natural language processing techniques, particularly transformer-based models, be combined with traditional search methods to

CEU eTD Collection

enhance the detection and prediction of greenwashing across global corporations, including but not limited to those listed in the Dow Jones Industrial Average? (DJIA)"

The **null hypothesis (H0)** of the research suggests that there is no significant effect of combined traditional search methods and NLP techniques in the detection and prediction of greenwashing practices among companies that are listed on the Dow Jones Industrial Average index. Statistically, this would mean that the results are due to pure chance and that the methods used in the research do not provide an improvement over existing greenwashing detection methods or over random guessing.

The alternative **hypothesis (H1)** suggests that the combination of traditional search methods and NLP techniques improves the detection and prediction of greenwashing practices among companies listed in the Dow Jones Industrial Average index. Moreover, that combining traditional search and NLP methods leads to statistically significant results and enhances greenwashing identification more effectively than do traditional or less advanced methods alone.

1.3 Thesis outline

The thesis starts with an introduction which defines the concept of greenwashing and its significance in the current corporate environment. The section also includes the research question and hypotheses. Following the introduction, the literature review explores existing research on greenwashing, and describes various greenwashing approaches. The methodologies used for greenwashing detection and their effectiveness are introduced. The literature review section also

CEU eTD Collection

highlights the gap in existing research and introduces how the thesis aims to address this via the proposed framework.

The methodology section describes the research design and the methodology used to develop the NLP-infused framework. This part provides detailed descriptions of how data was collected and processed, the introduction of the BERT model, and the development of custom definitions that were tailored to capture different greenwashing approaches more effectively. The model architecture section explains the technical architecture of the BERT model, including how traditional search methods are integrated with BERT in the research, to enhance greenwashing detection capabilities. The rationale behind choosing this particular model is also stated, as well as what are the expected benefits over other existing techniques.

These sections all build up to the results part of the research, which presents the outcomes of applying the framework to companies of the Dow Jones Industrial Average index. A quantitative analysis can be found here regarding the model's performance, with a special focus on precision, recall, and F1 score. How the results relate to the hypotheses is also discussed in the same section. In the discussion, the broader implications of the findings are explored. This section shows the practicality of the proposed framework in real-world settings such as regulatory practices and corporate behavior. Limitations of the current research are also addressed here, along with recommendations for future research to build upon this research and framework.

The conclusion summarizes the main insights of the research and concludes the thesis's contribution to greenwashing detection. It discusses the significance of these main contributions regarding academic research and practical applications in environmental and computer science.

CEU eTD Collection

2. Chapter: Literature Review

The literature review attempts to bring the phenomenon of greenwashing closer to the reader and examines the evolution of greenwashing detection methodologies. It begins with explaining why greenwashing is a threat currently, then proceeds with an overview of key definitions and types of greenwashing. The focus then shifts to the role of NLP, assessing existing literature on its effectiveness and efficiency compared to older and standard keyword-based approaches. The literature review highlights the most relatable and significant studies in the field, making conclusions about technological advances of greenwashing detection, and on the implications of these.

This research needs to be contextualized within the academic discourse, as greenwashing is a complex and multi-layered phenomenon. This section does just this, aand it also emphasizes the importance of innovative greenwashing detection solutions.

2.1 Greenwashing slows down sustainable progress

A side of greenwashing that to this day has received less attention is the lack of verified and available greenwashing data, which would encourage consumers to check where and what kind of products they are purchasing and think twice before they support a manipulative corporation. Unfortunately, the lack of unavailable data is hindering the progress of transforming to a more responsible way of consumption, especially in the U.S., where the consumption situation is worse compared to the rest of the world. Already 12 years ago, in 2012, David Tilford of the Sierra Club, the United States' oldest grassroots environmental organization reported the following:

CEU eTD Collection

“A child born in the United States will create thirteen times as much ecological damage over the course of his or her lifetime than a child born in Brazil”.

This research’s topic is strongly connected to Sustainable Development Goal 12, "Ensure sustainable consumption and production patterns" (“Responsible Consumption and Production” for short). According to the United Nations (United Nations 2024), since it undermines trust and deceives not only customers but the public and investors too with misleading claims or marketing, greenwashing is undermining genuine efforts towards genuine sustainability initiatives.

Since the 2015 Paris Agreement, at least nine thousand companies have joined the UN’s “Race to Zero Coalition”, setting net-zero goals. This includes half of the thousand top largest publicly traded companies (United Nations 2024). According to the UN definition, net-zero targets are those at which a company's residual emissions can be absorbed and stored by carbon sinks (such as forests and seas) alone. Although more and more companies join these efforts, their plans often lack transparency, and in most cases their focus is on carbon offset credits, which is undoubtedly different from changing their core operations and cutting real emissions (Offset Guide n.d.). Purchasing carbon credits from “under-emitter” companies is by no means equivalent to taking one ton of carbon dioxide out of the atmosphere. The clarity and trustworthiness of sustainable claims are thus extremely poor, according to the New Climate Institute and Carbon Market Watch’s 2023 Corporate Sustainability Monitor (Carbon Market Watch 2024).

According to the UN’s latest Sustainable Development Goals Report from 2023, sixty-two countries and the EU introduced 485 policies altogether to achieve the targets of SDG 12, however, deceptive environmental claims make it difficult to track progress (United Nations 2024). Without

CEU eTD Collection

verified data on a company's true environmental impact, it is challenging to measure success and identify areas for improvement.

At least six targets of SDG 12 are affected by greenwashing. By creating an illusion of sustainability commitment without actual sustainable core business values, greenwashing sabotages Target 12.1, which relates to implementing the 10-year sustainable consumption and production framework. This is because companies might greenwash their practices related to material sourcing or water usage, hiding their true environmental impact, thereby also impacting Target 12.2, which is about sustainable management and use of natural resources. Food production or packaging companies can mislead consumers about their efforts to minimize food waste throughout the supply chain, which impacts Target 12.3 - halve global per capita food waste. Without strict regulation, companies can overemphasize recycling initiatives or relevance of recycling in their business model, thereby contradicting Target 12.5, which aims to substantially reduce waste generation by 2030. Without a doubt, a form of greenwashing can be cherry-picking information about the environmental impact of a company - which distorts reporting, conflicting Target 12.6: encourage companies to adopt sustainable practices and sustainability reporting. Lastly, greenwashing hides the true environmental impact of goods and services, thereby lowering the awareness of consumers. This makes it difficult for them to adopt choices that promote sustainable living. Target 12.8, which is "promote universal understanding of sustainable lifestyles" is therefore also affected (United Nations 2023).

The importance of discussing and researching greenwashing is based on the fact that, as these examples had shown, of the eleven SDG 12 targets, greenwashing risks the achievement of at least six. The indicators of these six SDG 12 targets are essentially immeasurable unless legislation, corporate sustainability attitudes, or consumer behavior change, as businesses will

CEU eTD Collection

continue to find innovative methods to include greenwashing tactics in their marketing, reporting, and in their communication overall.

While it is most closely connected to SDG 12, greenwashing also connects to several other SDGs, including SDG 7: Affordable and Clean Energy through greenwashing the reliance on fossil fuels, and SDG 8: Decent Work and Economic Growth through exploiting workers in supply chains that are portrayed green. As discussed earlier, greenwashing can be prevalent in achieving or rather delaying climate action goals, which essentially form SDG 13. Industries that directly contribute to environmental damage such as deforestation or ocean plastic pollution, might engage in greenwashing to downplay their negative impacts, thereby halting SDG 14 and 15 - Life Below Water and Life on Land. Even SDG 16, Peace, Justice, and Strong Institutions is impacted by greenwashing, since resource extracting companies can use it to mask their activities and impacts in developing countries, and their contribution to social and political instability.

Altogether, greenwashing impacts at least seven out of the seventeen Sustainable Development Goals. Soergel et al. write about “co-benefits”, stating that even achieving the targets of one SDG could have a positive spillover effect on others (Soergel et al. 2021). Other authors, for example, Nilsson et al. confirm this, acknowledging that progress made on individual goals creates a momentum for progress on others (Nilsson, Griggs, Visbeck 2016). This was the main motivation to select greenwashing as the topic of this thesis. It can be expected that the elimination of the multi-layered greenwashing phenomenon in the current global sustainability landscape can lead to major advancement in global sustainability agendas.

CEU eTD Collection

2.2 *Classifying greenwashing: types and categories*

In the research "Concepts and forms of greenwashing: a systematic review" the authors thoroughly investigated the problem of greenwashing between 2010 and 2020, focusing on important classifications and typologies (de Freitas Netto et al. 2020). Their goals were to distinguish between the various types of greenwashing and its manifestations, and the authors identified claim and executional greenwashing classifications on firm-level and product-level. Claim greenwashing involves using written or implied arguments about the environmental advantages of a product or service (de Freitas Netto et al. 2020). The study cited research by Carlson et al. where the authors presented five typologies of green claims: product orientation, process orientation, image orientation, environmental fact, and combination (Carlson et al. 1993). The second classification type, executional greenwashing strategy is one that does not use any type of claim but uses nature-evoking elements such as images using the green color or bird sounds. These elements can subtly influence consumers' perceptions of a brand's greenness.

Firm-level claimed greenwashing refers to the misleading environmental practices of an organization or the environmental benefits claimed for a product or service at the company level. This type of greenwashing involves exaggerated or false claims about the environmental friendliness of the overall practices of a firm or its specific products and services. It can include claims such as promoting sustainable practices or products that are not representative of the company's actual environmental impact or activities. A fictional renewables investment company, for example, could be simultaneously lobbying against stricter environmental regulations, raising concerns about true commitment to sustainability. Lastly, firm-level executional greenwashing relates to a symbolic strategy whereby firms seek to gain or maintain legitimacy by disproportionately revealing beneficial or relatively sympathetic performance indicators to mask

CEU eTD Collection

their less impressive overall performance. The researchers also covered the idea of "decoupling behaviors" which refers to a situation where there is a disconnect between companies' symbolic and genuine environmental protection actions, and "greenwashing as selective disclosure," which occurs when businesses disclose positive environmental performance while withholding negative information.

Although greenwashing has been researched since 1986 (Orange and Cohen 2010), its systematic categorization is surprisingly recent and has not been identified until the publishing of the above research. Classification makes it possible to identify potential loopholes in regulations and standards that companies might exploit to make misleading claims. In the era of big data, it has been increasingly difficult for consumers to find quality information on the real environmental impact of a brand or product. Although information overload is thoroughly discussed below, it is crucial to mention data manipulation right in the introduction of this thesis. Fabricating data has become simple. Corporations can easily skew reporting, inflate numbers, or hide negative environmental impacts among the overwhelming amount of information they report, blurring the lines of transparency.

Tricking consumers into purchasing decisions based on false or exaggerated claims of sustainability and environmental friendliness leads to them unknowingly supporting and financing companies whose practices harm the environment (Visser and Hollender 2010). This misinformed choice is against consumer protection in the European Union (European Parliament 2024). In the US, however, where per-capita consumption and per-capita emission is peaking, there is little enforcement and most guidelines related to greenwashing are unclear or incomplete (Olatunji 2023). Shockingly, for example, no guidelines cover claims related to water use or packaging.

CEU eTD Collection

There is minimal regulation of greenwashing, even though consumption per capita increased by roughly 65% between 1990 and 2015, compared to a 35% growth in Europe (Schrager 2021).

Besides eroding trust of consumers, greenwashing enables companies to portray an environmentally friendly picture of themselves and continue environmentally harmful activities. Essentially, this delays the adoption of genuine sustainability practices and the achievement of collective environmental goals and SDGs (Montgomery, Lyon, and Barg 2023). The insights of the papers discussed above show that the method proposed by this research encompasses both claimed and executional greenwashing, emphasizing the need for a method to identify greenwashing practices that today manifest in increasingly more ways.

As Lyon and Montgomery make it clear in “The Means and End of Greenwash”, there is no concrete definition of greenwashing due to its multifaceted nature (Lyon and Montgomery 2015). The article also notes that there are various definitions of greenwashing mainly since it is discussed from different conceptual angles. As such, the term "greenwashing" therefore has no single, widely-accepted definition found in current literature. From the research's perspective, delving into different perspectives of greenwashing is crucial as different greenwashing types may have different results in terms of detection. Given the subjective nature of greenwashing and the absence of a standardized definition, it was essential to explore multiple viewpoints for a more comprehensive and less biased analysis. Although greenwashing is most often perceived and discussed as deceptive practices by firms regarding sustainability, such as misleading green claims, scholars often study it in terms of differing communication and action, selective disclosure, or as a method of decoupling, too. By deception or misinformation, we mean that a company engages in outright deception and spreads misinformation about their environmental practices. A company might label a product as "made with recycled materials," but when thoroughly inspected, only a

CEU eTD Collection

trivial amount of the product contains recycled materials. Such deceptive claims are unsupported by scientific data and not only mislead consumers but also undermine genuine efforts towards sustainability. An example for this type of greenwashing is the already discussed and well-known Volkswagen scandal. The company marketed diesel cars as eco-friendly with campaigns like “Think Blue”, however, they were caught having installed a software that cheated emissions tests, thereby making Volkswagen cars appear cleaner than they were. In reality, these cars emitted much more harmful pollutants (United States Environmental Protection Agency 2023).

Another approach of greenwashing is when the environmental responsibility claims of a company are inconsistent with its real actions. Companies may heavily market their environmental initiatives, such as recycling programs, but at the same time, neglect addressing significant issues like carbon emission reduction or waste management. The fast fashion industry is an example for this type of greenwashing, as many of the brands advertise “sustainable” collections or “eco-friendly” materials, when their overarching business model is based on high-volume production and low prices (Earth.org 2022). Their green initiatives are an insignificant fraction of their overall production.

Selective disclosure is a greenwashing strategy employed by companies to highlight certain positive environmental aspects while conveniently omitting negative ones. For instance, a company might boast about its efforts in water conservation but fail to mention its substantial energy consumption or carbon emissions. By cherry-picking and highlighting the good details while downplaying less admirable ones, these companies show an incomplete and skewed image of their environmental practices, which ultimately misleads consumers and even business stakeholders. Without mentioning names, many companies promote recycling initiatives, but rarely address the significant plastic waste they produce globally. A popular soft drink

CEU eTD Collection

manufacturer produces around 500 billion bottles every year - visualizing this means 150 million tennis courts filled with these (The Guardian 2019). This company is not the only example.

Lastly, there are scholars who see greenwashing as a CSR decoupling effort, for example, Yang et al. (Yang et al. 2020). They approach greenwashing as the company's brand built around a green image which deliberately hides actual environmental impact of operations. Let us imagine a company that sells bottled water and claims its bottles are made from recycled plastic that was collected during cleanups. Its marketing strongly emphasizes initiatives like these, but the company never mentions water sources or the environmental impact of large-scale bottled water production (especially compared to tap water). This is decoupling their eco-friendly marketing from their core product's environmental impact. Oil companies often exploit this, there are companies with programs that might appear environmentally conscious, but these programs do not reduce their own emissions, and do not decouple their image from their core business practices. While decoupling and the misleading communication in contrast to action greenwashing types are similar, there is one key difference - the intention behind this behavior (Bothello et al. 2023). Miscommunication about ongoing CSR efforts can in some rare cases stem from oversight, branding an entire company around these initiatives suggests an intentional strategy (Pedersen et al. 2024).

It is crucial to highlight that while differences between different greenwashing types might seem small, their different results are significant. As historical examples have shown, the consequences for a car manufacturer caught cheating emissions tests will be far greater than a clothing company's misleading "sustainable collection." Among other things, public outrage and government fines would differ significantly. Besides, some greenwashing tactics are more easily exposed than others - Volkswagen's scandal involved provable evidence, leading to a huge

CEU eTD Collection

publicity. The specific outcomes of greenwashing depend on unique circumstances of each greenwashing case, this is why it is becoming increasingly difficult to find an ideal and generalizable greenwashing detection method.

2.2.1 Review of greenwashing approaches in academic literature

The Merriam-Webster and Cambridge dictionaries describe greenwashing as a company making something environmentally friendly when it is not (Merriam Webster Dictionary 2024; Cambridge Dictionary 2024). De Freitas Netto et al. in their 2020 research also grasps the essence of greenwashing from this direction (De Freitas et al. 2020). The combination of Oxford dictionary definitions, however, places emphasis on the misleading nature of greenwashing (Oxford Dictionary 2024). According to UL Solutions (until 2007, Terrachoice), greenwashing is when companies selectively disclose information, and focus on positive environmental aspects while omitting negative ones (UL Solutions 2024). In other publications, writers like Tang et al. (2023) and Talpur, Nadeen, and Roberts (2023) describe greenwashing as a decoupling activity in which businesses make symbolic gestures to look environmentally conscious but do not genuinely act. As these examples demonstrate, there are several ways greenwashing can be explained, but there is no one, standard definition.

The examination of literature on greenwashing reveals at least the following greenwashing approaches: deception and misinformation, misleading communication in contrast to action, selective disclosure, and greenwashing as decoupling. These are also the approaches that were used when creating “custom”, all-encompassing definitions for the code of the model.

CEU eTD Collection

By incorporating multiple perspectives from the literature review into the analysis, this research avoids locking itself in a certain section of greenwashing practices, or in using an arbitrarily selected greenwashing definition. Working with one definition could introduce bias, especially considering the quantitative nature of the analysis focusing on sentence similarity. Covering more views and perspectives allows for a broader examination of the examined articles on greenwashing. The results section of the research demonstrates in more detail why considering these four viewpoints can be considered as an efficiency enhancement of the model.

2.3 The evolution of corporate social responsibility

This section sheds light on how societal expectations have shaped businesses' roles in addressing social issues throughout history, serving as a crucial theoretical base for the research. Roman laws and societal structures that were founded to solve social issues are the origins of CSR (Atkins and Osborne 2006). Through the Middle Ages, “businesses” at that time were viewed as tools for social progress, which brought about charity and social welfare facilities (asylums, hospitals, and orphanages) (Davoudi et al. 2018). During the Victorian era, social reforms, mainly humanism, gained popularity.

Over the course of the 20th century, CSR saw a transformation as it evolved from a vague idea to one that is widely applied by businesses. Academic discourse on the social performance of businesses already emerged in the 1930s and 1940s, but the early phases of we call CSR today were shaped by academic literature in the 1950s and 60s, with Bowen (1953) defining the specific social responsibilities of business executives. In 1960, Keith Davis emphasized the social responsibility of businesspeople and argued that they so have obligations towards society in terms of economic and human values (Keith Davis 1960).

CEU eTD Collection

In the 1970s, awareness of the rights to environmental, worker, and human protection increased even more. Government agencies were soon established, including the U.S. Environmental Protection Agency in 1970, the Occupational Safety and Health Administration in 1971, as well as the U.S. Consumer Product Safety Commission in 1972. These agencies all acknowledged that social relations between industry and society were changing, evolving.

A strategic approach to CSR emerged in the 1980s, with a particular emphasis on putting CSR into operation and executing it well. Some notable examples from this decade include articles by Thomas M. Jones (1980), Tuzzolino and Armandi (1981), Strand (1983), and Cochran and Wood (1984). These articles all went beyond just explaining the concept and offered methods for evaluating CSR from an operational perspective. This time was influenced by global trends related to sustainable development and a growing awareness of environmental protection. The Montreal Protocol (1987) and the establishment of the Intergovernmental Panel on Climate Change in 1988 are just two examples (Agudelo et al. 2019).

CSR became institutionalized in the 1990s, Burke and Logsdon presenting the concept of strategic corporate social responsibility (SCSR) in 1996 and emphasizing the link between CSR and its positive financial performance of the firm (Burke and Logsdon 1996). Elkington created the idea of the "triple bottom line" in 1994 as a sustainability framework that considers the economic, social, and environmental effects of a business (Elkington 1998). Donna J. Wood (1991) developed a model of Corporate Social Performance (CSP) that defined three dimensions - principles of corporate social responsibility, processes of corporate social responsiveness, and outcomes of corporate behavior (Wood 1991). Her model was more comprehensive and wide-ranging than those of Carroll (1979) and Wartick and Cochran (1985), and it explicitly highlighted the potential positive financial benefits of implementing CSR elements in a business framework.

CEU eTD Collection

The creation of the United Nations Global Compact in 2000 elevated CSR to a global level, as its primary objective was to inspire companies worldwide to adopt socially conscious and sustainable practices (Kadyan 2016). It also required companies to report on how these practices progress. The above situation established a "social contract" that required corporations to play a significant part in achieving the 17 Sustainable Development Goals.

Throughout the 2000s and beyond, CSR recognition grew, focusing on human and labor rights and sustainable development. Companies are now being urged (both by stakeholders and by regulators) to address social and environmental issues. However, CSR is often and more increasingly used as a strategy for value creation, which brings about a threat in the form of greenwashing practices.

2.3.1 Clarifying corporate social responsibility, corporate sustainability, and ESG

The term “corporate sustainability” is frequently used by businesses that aim to implement or scale up a sustainability related philanthropy (Fischer et al 2023). However, it is being almost interchangeably used with the term “corporate social responsibility” (CSR), when in fact, corporate sustainability more accurately falls under “environmental, social, governance” (ESG). Knowing the difference between the three terms is crucial to be discussed in the research because blending definitions can lead to the misinterpretation of a company’s sustainability statements from all stakeholders’ side. The commitment of a company to long-term environmental, social, and economic sustainability is corporate sustainability (Beattie 2024). Its objective is to create long-term benefits for all stakeholders while protecting people, the environment, and the economy. These three are also the three pillars of ESG, an increasingly widespread framework that focuses on assessing the risks of business related to sustainability (6clicks 2024).

CEU eTD Collection

Of the three pillars of business sustainability, the environmental pillar is the subject of the greatest discussion. It covers the several steps businesses may take to lessen their carbon footprint and environmental effect; reducing the amount of packaging waste and recycling materials are just two examples. As per the social pillar, the company's efforts to get the approval of stakeholders, staff, and the society at large are the main emphasis (Tenney 2024). Giving back to society, lobbying against child labor and exploitation of workers are examples of social pillar values. Implementing sustainable business practices to support long-term profitability is the third, economic pillar's goal. After all, the less lucrative a business is, the less beneficial it can be for the community or the environment. Excellent corporate governance and compliance are components of the economic pillar, meaning that when it comes to resource allocation, management and stakeholders aim to share the same ideals. If the economic pillar is strong, a corporation may invest in and develop new corporate sustainability strategies.

Having stated this, no pillar should take precedence over the others. In that case, it might happen that companies are exposed to seeking profit and taking short cuts - for example, by making greenwashing claims thereby incorporating deceptive manipulation into their messaging. Building on corporate sustainability and ESG, CSR involves voluntary actions taken by a company to address social, environmental, and economic issues, going beyond legal requirements. CSR has many elements, and not every firm has the resources to focus on them all. Some examples can include volunteerism, ethical sourcing, and social programs. Making investments in every project that comes up is not the aim. CSR initiatives must have a focus on fostering trust and creating a sense of connection with customers, at the same time, encourage companies to be more transparent about their operations and accountable to their stakeholders (Yeyi et al. 2023). Since the focus of this thesis is partly on sentence similarity analysis, understanding how these

CEU eTD Collection

terms are used in news articles is crucial for accurately interpreting information and effectively detecting greenwashing.

2.4 Greenwashing and the media

The above mentioned holistic corporate social responsibility approach is exemplified in real-world cases like Volkswagen's manipulation of emission tests, where the breach of trust not only harmed its reputation and short-term revenues but also underscored the crucial need for authenticity in corporate sustainability communication.

Trying to portray a green image, Volkswagen manipulated emission tests of 11 million cars between 2008 and 2015 (Jung and Park 2017). From the perspective of all market participants, including customers, this was a breach of trust. The scandal brought to light the fundamental problems with the authenticity of corporate sustainability communication. The “Dieselgate” scandal is the biggest to date in the automotive industry, and for many, it is also the most remarkable greenwashing scandal (Davison 2024). In the Volkswagen case, the intentional manipulation added a level of outrage to the situation. As more and more businesses adopt sustainable practices, cases like this emphasize the need for methods that enable customers and regulators to detect deceptive corporate behavior early on (as a reminder, Volkswagen had been alternating emissions data for six years before their cheating software was discovered).

Public impressions about corporate greenwashing are greatly influenced by news outlets, as the Volkswagen scandal shows. Media platforms in the 21st century function as influencing channels for spreading information, whether true or false, and this impacts the consciousness of consumers about global issues (OECD, 2022). Many authors have found a correlation between

CEU eTD Collection

media sentiment on sustainability practices and corporate financial performance. To name a few studies, Akyildirim et al (2023) shed light on the relationship between negative media narratives, greenwashing, and subsequent fluctuations in stock prices. Similarly, Strycharz, Strauss, and Trilling (2018) find connection between media communication and stakeholder attitudes towards corporate sustainability. This is consistent with Tetlock's earlier, 2007 research, where he found that media narratives have the power to influence stock prices in addition to public perception. To this, the proliferation of greenwashing activities over the last 15 years added another, CSR related dimension.

The media is a narrative shaper which is evident in various forms, from news coverage to investigative journalism (Karadimitriou et al. 2022). An illustration of this can be observed in the aftermath of the BP Deepwater Horizon oil spill. Kassinis and Panayiotou agree in their 20017 research that BP's 200-million-dollar investment in green advertising mitigated the reputational damage of the oil spill disaster, raising important questions about the incentive for firms to greenwash. From this example, while in general, the media plays a role in exposing greenwashing firms, there is a need for a clear definition of greenwashing, in order not to skew our understanding of individual cases. Until it can be openly defined by corporations, the media, and law, corporations can and will exploit this ambiguity significantly, and get away with serious environmental cases.

While the media plays a key role in uncovering greenwashing, at least to create awareness, a standardized definition of greenwashing would be required to avoid misunderstandings and hold firms accountable for their sustainability efforts. Until the issue of this common definition is resolved, NLP-based models such as the one in this thesis can be a useful tool for detecting greenwashing.

CEU eTD Collection

2.5 *Greenwashing affects consumer psychology*

According to Sun and Shi (2022), when consumers feel that a company has betrayed them by engaging in greenwashing, it further reduces their willingness to purchase green products. Anne Brouwer collects three impacts of greenwashing on consumers in her 2016 research - perception, purchase intention, and purchase behavior, finding that most people are surprised, disappointed, or even shocked when a greenwashing company gets uncovered. However, her study finds that in most cases, the disclosure of greenwashing practices does not lead to changes in consumers' perception of the company's credibility and brand trustworthiness.

Another thing Brouwer finds is that varying impact greenwashing has on purchase intent of consumers. The results of the group discussion suggest that only some participants are reluctant to continue purchasing the products of a company after a greenwashing news, while most people indicated that the information did not influence their buying behavior. This aligns with the findings of Smith and Brower from 2012 and Urbanski and Adnan from 2020. Based on survey responses, Smith and Brower found that there is a trend of consumers continuing to buy products with false labeling such as "biodegradable", even when they are aware that the firm has been involved in a greenwashing scandal before. This study was published 12 years ago, the greenwashing landscape has only gotten worse since, due to the current information overload, discussed in detail below. Based on the answers of 768 participants, Urbanski and Adnan found that even though consumers were doubtful, they still perceived greenwashed products as being good for the environment.

Furthermore, according to Brouwer, when making a purchase, people still tend to evaluate brand, price, and quality more than environmental factors or eco-friendly claims like reusability.

CEU eTD Collection

From this CSR efforts can be obscured by greenwashing, leading consumers to prioritize more familiar buying factors.

After Brouwer's group discussions, several participants expressed a desire for more information about greenwashing to make better-informed purchase decisions. Brouwer has found that most people who continue to purchase goods from businesses that participate in greenwashing do so because they find it hard to find transparent environmental data. The creation of a centralized database for corporate environmental data would be a significant advancement, and this is where this research can help. The collective data could include data on a company's environmental impact reports, their sourcing practices, supply chains, but also any documented greenwashing accusations. To tackle greenwashing, the empowerment of consumers through accessible and transparent environmental data is a crucial step. Only by doing this can those market segments thrive which reward genuine sustainability efforts of firms. As it will be described in the "Limitations and implications for future research" part, the method proposed by this thesis could be further improved and extended, to cover all news outlets and firms on a global scale, thereby fostering a culture of transparency.

2.6 Information overload

In today's world, consumers are faced with the challenge of shifting through news articles and reports to get to the relevant information. Across industries, there is no universally accepted guide to expose greenwashing for consumers. Only the most dedicated people conduct thorough research before making a purchase. Moreover, a survey by Blue Yonder in 2023 revealed that only 56% have doubts about the sustainability claims made by brands (Blue Yonder 2023). This discrepancy suggests that consumers are not actively verifying claims. Navigating complex environmental data

CEU eTD Collection

and feeling overwhelmed by the sheer amount of information available online is not an uncommon experience. It is difficult to find readily available, comprehensive sustainability guides that consumers can use to identify greenwashing. While conducting this research, it quickly became apparent that the most comprehensive and most up-to-date greenwashing guides available online are independent websites like the GreenWashing Index, Changing Markets Foundation, and Greenwash. They offer valuable resources, but even they are not much different from simple news collections about greenwashing. Unfortunately, their reach is not extensive enough to get to all consumers, and their specific focus on certain sectors limits their applicability. Overall, the lack of readily available and trustworthy resources presents a challenge for consumers in informed decision-making.

Even with quality resources available, evaluating sustainability claims of companies can be difficult, because they often use complex jargon or vague language. This makes it challenging to evaluate the actual environmental footprint of the products or services they use.

WWF recommends checking four things - buzzwords, evidence, verification, and sustainability (WWF 2024). However, most buzzwords like “green” and “environmentally friendly” have no legal background, these mean nothing without additional evidence of data that can be verified. Recognizing the limited availability of relevant environmental information online, and the danger of fabricated data, the thesis aims to simplify the process of finding information related to a company’s greenwashing activities. By combining traditional analysis with advanced language model sentence similarity methods, the goal is to streamline the evaluation process of consumers and introduce a method which is not only accurate but also relevant to assessing greenwashing activities.

CEU eTD Collection

In 1948, during the Scientific Information Conference of the Royal Society, overload was officially acknowledged (Bawden and Robinson 2020). Although the phrase "information overload" had not yet been created, the vast number of potentially relevant research material that was circulating generated concerns that scientists might become so overwhelmed that they would lose productivity and possibly become less professional.

When someone feels overwhelmed by the amount, diversity, and complexity of information available to them, they are said to be experiencing information overload. When it comes to a person's capacity to use information effectively and efficiently for their work, education, or just everyday living, it can be disturbing and intrusive. It usually leads to feeling overwhelmed, feeling a loss of control, and anxiety. It is especially harmful because it may impact a person’s ability to make good decisions. Besides consumers, there are other stakeholders who can be impacted by information overload when trying to assess greenwashing practices of companies - investors, regulators, environmental advocacy groups, and of course, researchers and academics. Therefore, it is essential to comprehend and manage information overload to effectively evaluate the environmental performance and obligations of companies.

While until the 90s, information sources were mostly books and journals, these are now mostly digital sources such as social media and online news. Information overload has been a long-standing problem which has only gotten worse. One of the biggest challenges in managing information effectively is the fact that it inherently needs to be selectively filtered. Some information must inevitably be excluded throughout the process, but because there is no way to confidently know what is unnecessary, there is a chance that important insights might be overlooked. It is important to stress that one of the fundamental issues that all stakeholders have with greenwashing is the deluge of information available on the internet, paired with the lack of

CEU eTD Collection

credible responses about how businesses are affecting the environment, especially when it comes to product-level information. Consequently, only the most dedicated ones are willing to invest time in researching before making purchases or supporting companies financially.

Numerical data is important in illustrating the volume of digital information that consumers in the twenty-first century must deal with. Bawden and Robinson (2009) assert that one weekly news magazine today has access to more information than a person living in the 17th century would in their lifetime. In 2013, SINTEF data showed that ninety percent of the world's data was generated within the years 2011 and 2012.

The internet itself is growing at a 26% rate per year, in fact, in 2017 alone more data was created than in the previous 5,000 years of humankind (Rao, 2023). However, these numbers sadly only mean volume, not quality or value. In a TMMData poll with 800 qualified answers, just 28% of respondents said they are "extremely/very clear" about the source of their marketing data (The Data Warehousing Institute 2017).

In the digital age, information is a double-edged sword for consumers. The volume offers exceptional access to knowledge, but it creates an information overload. This, coupled with the scarcity of relevant, verified answers concerning environmental impact makes environmentally deceptive activities of companies thrive. Despite the expansion of digital sources, the need for reliable answers persists. This gap has accelerated the exploration of AI-driven solutions to detect and mitigate greenwashing effectively across industries.

CEU eTD Collection

3. Chapter: Methodology

3.1 Introduction

To comprehend the intricate and intertwined connection between corporate sustainability initiatives, public attitudes, and economic results, a comprehensive approach is necessary. This section introduces the research's methodology, which aims to investigate these connections, with a particular emphasis on the integration of traditional search methods with NLP techniques. The methodology is structured around the proposed framework that combines traditional keyword search with AI models, particularly focusing on Bidirectional Encoder Representations from Transformers (BERT). The aim is to capture the nuances, sentiment, and contextuality in media content in a way that these can be quantitatively analyzed later in the results section of the research. The methodology's core is the concept of sentence similarity, which allows for a deep understanding of the semantic relationships between news articles and our predefined criteria related to greenwashing. The methodology tackles the difficulty of handling overwhelming amounts of information when analyzing extensive datasets. The following sections will provide a detailed explanation of each component of the methodology.

3.2 Existing NLP techniques for greenwashing detection

Researchers have been interested in assessing the difference within sample sets since the 1880s (Murphy 1996). The first model to measure difference between sequences was the Levenshtein distance in 1965, and it is still a major element of computer science (Levenshtein 1966). It tells the user how different two strings (consequence of characters) are. A simple example would be the Levenshtein distance between "rain" and "shine," which is 2. This number ("distance") symbolizes the smallest number of edits needed to change one word into the other.

CEU eTD Collection

After these lexical and syntactic approaches came knowledge-based models. WordNet is the most known example for these. It is a path-based similarity measure from the 1980s, containing the largest database of the English language, grouped into sets of cognitive synonyms (Fellbaum 2005). Semantics and lexical relations connect these. The result is a network of (meaningfully) connected concepts and phrases, which can be used for natural language processing. Later, corpus-based approaches became popular, which means models that conduct statistical analysis on vast collections of written or spoken data (corpora).

The method closely related to the subject of this research is latent semantic analysis, which is also corpus-based. It suggests that words with similar meanings will appear in similar contexts in text, according to the distributional hypothesis (Harris 1954). Topic models also originate from this era, having been first defined by Papadimitriou et al. in 1998, and their results, clusters of related phrases, are still utilized in the media and in NLP research today (Papadimitriou et al. 1998). When a visitor of a news website searches for "tech" or "science" keywords, he is indirectly using a topic model.

The greatest progress in the study of sentence similarity was made in the new millennium. More advanced NLP techniques such as word embedding surfaced first around 2000 (Jurafsky and James 2000). For this research, a word embedding model was selected, namely, BERT (Bidirectional Encoder Representations from Transformers). However, it has also been refined with Siamese network theory and transformer-based models - which are the latest approaches to sentence similarity and actively facilitate deep learning (Huggingface 2024).

As mentioned in the first paragraph in the methodology section, this research leverages sentence similarity, an NLP technique. Sentence similarity allows for capturing the semantic similarity of news articles about pre-defined companies, providing insights into their potential use

CEU eTD Collection

of deceptive environmental messaging. Sun et al, propose a novel approach to sentence similarity. Their framework measures it by examining the likelihood of producing two sentences in a similar context.

This unsupervised method (meaning that the computer learns from unlabeled data) emphasizes the importance of context in understanding sentence meaning (Sun et al. 2022). There are key takeaways in this work. First, the authors argue that a sentence's meaning is best understood if we consider its surrounding context. Second, the authors demonstrate that comparing the probabilities of generating two sentences within the same context allows for quantifying semantic similarity. (This approach differs from traditional methods like bag-of-words or deep learning models, which focus on word frequencies or data representation rather than contextual meaning.)

Building upon the work of Sun et al., this thesis leverages a modified version of their sentence similarity framework to investigate greenwashing by companies. Their core idea of analyzing contextual meaning is particularly inspiring. Their framework is adapted in this research to compare the semantic similarity of news article titles related to the companies under investigation. By focusing on contextual meaning, this approach helps identify deceptive environmental messaging in these titles. This allows to explore articles potentially linked to greenwashing and assign "greenwashing probability scores" to the companies based on the analyzed news articles.

CEU eTD Collection

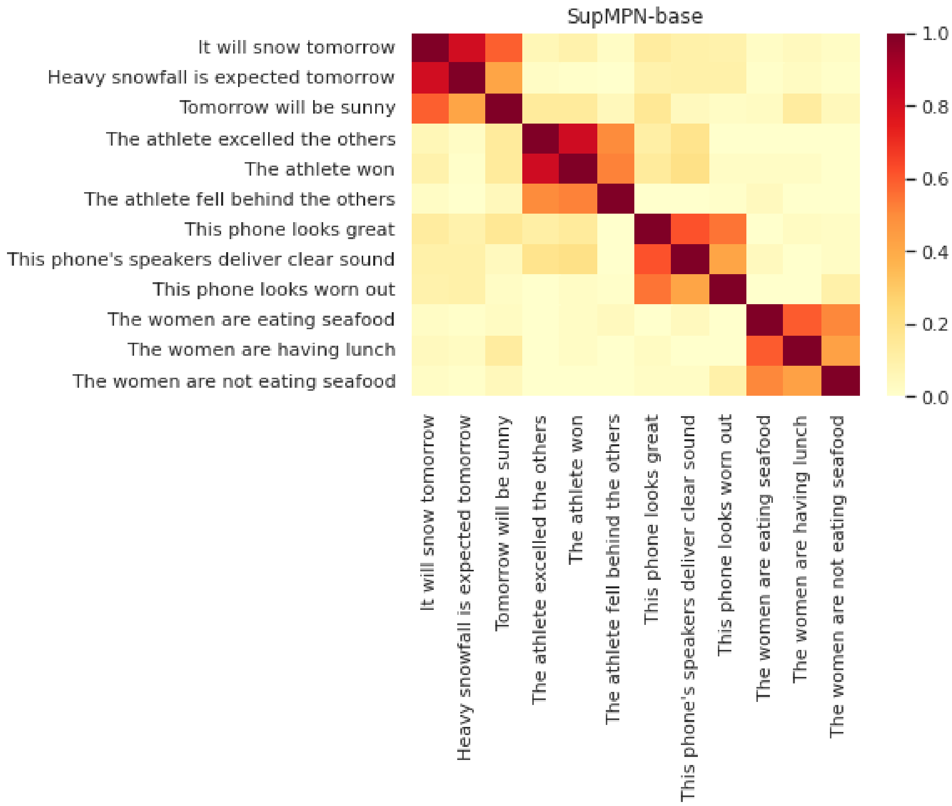

Figure 1: Illustration of how sentence similarity is measured, adapted from Dehghan and Amasyali (2022), 'SupMPN: Supervised Multiple Positives and Negatives Contrastive Learning Model for Semantic Textual Similarity'

Figure 1, taken from Dehghan and Amasyali's 2022 work gives an idea of how sentence similarity is measured (Dehghan and Amasyali 2022). Taking the "Tomorrow will be sunny" from the Y axis, and "Heavy snowfall is expected tomorrow" from the X axis as examples, their cross-section (their similarity score) lies somewhere around 0.4. This is not a pronounced similarity because one sentence relates to sunny weather, the other to snowfall - however, there is still a certain level of similarity as both sentences refer to tomorrow's weather.

CEU eTD Collection

Essentially, this is how the model of this thesis works - the code compares the search term and the title of news articles, then gives a probability score based on the similarity of these two. More on this is detailed in the model architecture section of the research.

3.3 Custom greenwashing definitions

The engine of my analysis is a Python code which attempts to determine if a company has been the subject of online articles discussing their involvement in greenwashing practices. While the code will be elaborated upon later, it is important to highlight that by introducing this model, the thesis offers a new way of finding accurate greenwashing related information online. It streamlines the desktop research process, and at the same time, handles biases, thereby saving valuable time for the subsequent data analysis.

What makes the model even more unique is that more approaches towards greenwashing are considered - deception and misinformation, misleading communication in contrast to action, selective disclosure, and greenwashing as decoupling. After reviewing about one hundred articles on greenwashing, these four types of greenwashing have been identified in literature, which serve as a key part of the research’s methodology.

As mentioned before, greenwashing has no standard, legal definition as of today. Rather than relying on arbitrary definitions, individual definitions were constructed for each of the four approaches within the model's code. These definitions, based on literature, form the foundation for comparing them to Google News article titles. It is important to note that although they do not need a specific terminology, transformers like BERT do need specific language and vocabulary. They need grammatically correct and easy-to-read sentences, words that are widely used (so there

CEU eTD Collection

is a bigger possibility of them having been trained on these and actually recognize them). The surrounding words and clear sentence structure is also important for understanding the meaning of a word (Hugging Face 2023). These points were taken into consideration when compiling the below definitions. Here, introducing some creativity was crucial to tailor definitions precisely to BERT, so even though all definitions were partly inspired by certain articles, they reflect significant original contribution.

<u>The four constructed definitions are the following (they are written here exactly as they are written within the Python code):</u>

- Deception and misinformation. "This is when companies engage in deceptive practices by advertising products as "eco-friendly" or "sustainable" without providing supporting evidence. By doing so, they mislead consumers and in the longer run also compromise genuine sustainability efforts. Their claims involve a range of misleading actions, including misrepresenting product ingredients or origins." (mainly inspired by De Freitas Netto et al. 2020).
- Misleading Communication: "Companies highlight trivial environmental initiatives, such as recycling programs or energy-efficient manufacturing, but at the same time, they downplay or ignore significant environmental impacts like emissions and pollution. This is a facade of environmental responsibility which misleads stakeholders about the company's actual ecological footprint. The strategy is deceptive by disproportion rather than fabrication." (mainly inspired by Visser and Hollender 2010).
- Selective Disclosure: "Companies selectively disclose positive environmental activities, cherry-picking data to present an overly positive view of their environmental impact. An example to this is highlighting achievements like participation in renewable energy

CEU eTD Collection

initiatives while omitting information about activities that have a significant negative environmental impact (such as elevated levels of waste production or biodiversity loss). Companies shift the narrative and mislead stakeholders by omission by presenting a distorted image that hides the actual environmental cost of their operations.” (mainly inspired by UL Solutions 2024).

- Greenwashing as Decoupling: “Companies create a mask of being sustainable through branding and symbolic actions that are decoupled from their actual environmental impact. These strategies often involve initiatives that - while visually appealing or publicly engaging -, do not result in substantial improvements to the company's environmental performance. Examples include high-profile but limited scope sustainability campaigns or using "green" imagery in marketing materials. This misleads consumers by creating a misleading impression of environmental responsibility.” (mainly inspired by Talpur, Nadeen, and Robers 2023 and Pedersen et al. 2024).

The “general” greenwashing definition was created by combining the dictionary meanings of greenwashing from the Oxford, Cambridge, and Merriam-Webster dictionaries:

"Greenwashing is the deceptive act of employing behaviors, activities, or communication strategies that deliberately create an inflated perception of a company's environmental commitment and performance. This manipulation can involve unsubstantiated claims of sustainability, misleading marketing tactics, selective disclosure of environmental data, implementation of superficial eco-friendly practices, and misrepresentation of a product's environmental impact." (by combining the Oxford English Dictionary 2024, Merriam-Webster 2024, and Cambridge Dictionary 2024 definitions).

CEU eTD Collection

3.4 *Transformer methods*

In 2017, artificial researchers of Google published “Attention Is All You Need”, research that explains the at that time novel Transformer architecture, which was pivotal for machine learning advancement, especially in natural language processing (Vaswani et al. 2017). The "self-attention" mechanism in this design replaces conventional recurrent layers, enabling the model to assess the relative importance of words regardless of how far apart they occur in a sentence. Recurrent neural networks (RNNs) process information sequentially, analyzing one element at a time (DiPietro and Hager 2020). For tasks like machine translation or sentiment analysis (where understanding the context of the entire sentence is crucial), this can mean limitation. The self-attention mechanism breaks free from this sequential processing. Imagine a student writing an essay. With an RNN, the student would have to write the introduction, then the first body paragraph, and so on, without being able to revisit earlier sections. Self-attention allows the student to consider all the ideas in the essay simultaneously while writing each sentence.

Self-attention enables the model of this research to assess the relative importance of words regardless of their position in the sentence. Therefore, it can more efficiently assess and comprehend the overall context of environmental claims. It does so by utilizing the transformers’ capacity to process entire sequences of data simultaneously rather than sequentially. This allows the model to spot nuances and potential discrepancies that may be signs of greenwashing. Transformers have essential characteristics, such as their self-attention structure, which offer a strong system for dealing with intricate data patterns, which is a critical aspect in differentiating between authentic environmentally efforts and greenwashing.

CEU eTD Collection

3.5 *Bidirectional Encoder Representations from Transformers (BERT)*

First described in 2019, BERT improved performance on a variety of NLP tasks such as the one used in this thesis, sentence classification (Reimers and Gurevych 2019). BERT leverages its Transformer network to generate semantic representations (sentence embeddings) of input sentences. These embeddings are fixed-size vectors obtained through a two-step process, Transformer Network Processing and Siamese and Triplet Network Structure. BERT utilizes its Transformer architecture, a deep learning model adept at capturing relationships between elements in a sequence. Within the Transformer, the BERT analyzes the sentence by dissecting individual words and their interactions with each other. This analysis can capture not only the explicit meaning of words but also the context they create within the sentence. The specifics of the Siamese and Triplet network structures are beyond the scope of this research, but it is nevertheless important to understand their role. These structures are typically used for learning similarity metrics - in BERT's case, they facilitate the conversion of the complex internal representation generated by the Transformer into a fixed-size vector, which serves as the sentence embedding. These embeddings are not simply random numerical codes, they are designed to be semantically meaningful, encoding the actual meaning of the sentence they represent. Sentences with similar meanings will have similar embeddings, residing in closer proximity within the vector space. This semantic nature allows for tasks based on the meaning conveyed by the sentences. An example for this can be these two sentences where sentence on is "This is an eco-friendly product.", and sentence two is "This product is kind to the environment." Both sentences have a similar message. BERT would analyze them, create sentence embeddings, and likely find a high cosine similarity between the two codes. This indicates their close semantic meaning.

CEU eTD Collection

The research uses BERT above other natural language models because it is particularly strong in not only understanding words and sentences but also the contextual meaning of those excels at understanding the contextual meaning of words and sentences. When analyzing articles about the environmental claims of firms, BERT can help in recognizing and understanding nuances in the language used, providing a more accurate understanding of the content. This is crucial when analyzing a sarcastic article title, for example. In Figure 2, it can be seen BERT can decide whether the word “it” refers to “animal” or “street”. In this example, it knows (self-attentively weighs) that “animal” is the correct word because this association has the highest calculated score out of all listed words.

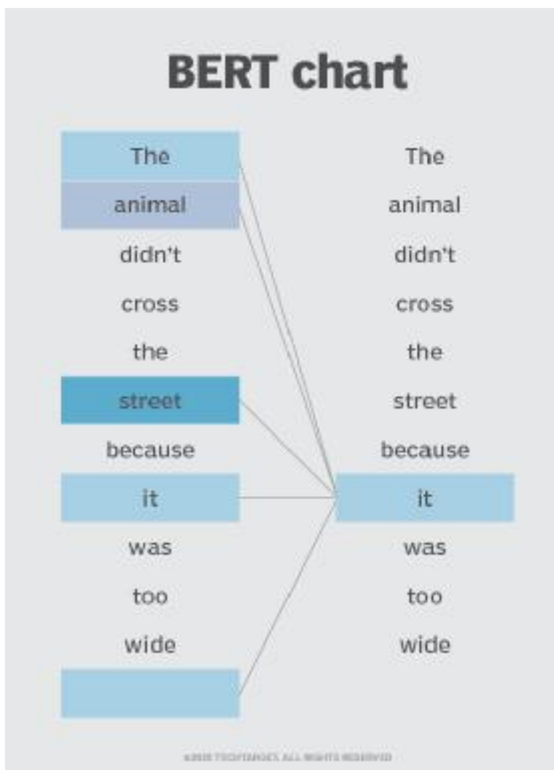

Figure 2: BERT's self-attention mechanism in action, adapted from Hashemi-Pour and Lutkevich (2024), 'BERT language model'

CEU eTD Collection

In conclusion, BERT can be an effective tool for greenwashing detection because of its capacity to capture both the explicit meaning of words and their context. BERT analyzes sentence structure and word relationships and generates meaningful numerical embeddings. This allows for identifying articles with similar environmental messaging, even if they use different wording.

3.5.1 Rationale for using an open-source model

The code, which is described in the methodology part of the research, utilized the "all-miniLM-L6-v2" pre-trained model to analyze a list of selected companies and their potential involvement in greenwashing scandals. This HuggingFace model was trained on 1 billion sentence pairs and can map sentences on a 384-dimensional vector space, which is about one fifth of the company behind ChatGPT, OpenAI's version of BERT, "text-embedding-ada-002" (OpenAI 2022). The reason why OpenAI's model was not employed in this research is because it is closed source and cannot be downloaded to local computers, which limits the research's reproducibility.

Open-source software are those for which the code is made available and can be shared and modified by anyone (Open Source Initiative 2024). Users have the freedom to modify these according to their needs, and certain licenses, for example, Apache, ensure that this freedom is maintained (Raymond 1999). In the case of closed source software, on the other hand, the source code is not made available to the public, as modifications and redistribution of closed source software are usually restricted by licensing agreements. The question of using open versus closed source is important for this research on greenwashing detection for several reasons. In the end, BERT was selected, which is open source. The transparency of open-source software enabled the understanding of how it operates, which ensured the credibility of the results obtained through its

CEU eTD Collection

use. Additionally, thanks to its open source, anyone can reproduce this research and verify the findings or further develop the model (McMillan and McLure 2003).

To detect greenwashing, BERT needed to be modified to assess contextual data about environmental sustainability and corporate communications. BERT's open-source nature allowed for tweaking the code and fine-tuning the algorithm to reach best performance and end results. In the methodology, BERT plays a crucial role and allows for comparing the pre-defined definitions of greenwashing with the headlines of news articles found on Google News.

3.6 Cosine similarity enhancement

As discussed earlier in the research, greenwashing claims today are communicated through various online channels such as corporate websites, press releases, several types of annual reports, and marketing materials. Text similarity methods can help analyze the similarity between these communications and known instances of greenwashing, which helps identifying those patterns or language cues that are associated with greenwashing practices (Cann et al. 2023).

There are several text similarity algorithms available, including the Jaccard similarity. Developed in 1884 by Grove Karl Gilbert, it determines the similarity by computing the ratio of the size of the intersection to the size of the union (Murphy 1996). It is commonly used for comparing the similarity between documents, based on their word overlap. There is another popular method, the Levenshtein distance, which calculates the number of edits required to change one word into another. Levenstein distance focuses on edit operations and is suitable for texts, while Jaccard similarity focuses on set intersection and is more applicable for sets of elements.

CEU eTD Collection

This study employs a third type of text similarity approach, Cosine similarity. It calculates the cosine of the angle (formed by two vectors) in a 3D space (Miesle 2023). It is commonly used in text mining and natural language processing to assess the similarity between documents represented as vectors. Compared to the Jaccard similarity or the Levenshtein distance, it is better suited for comparing texts with different lengths and contents since it is resilient to variations in document length and word frequency (Kang and Kim 2023). The present methodology's application of cosine similarity is consistent with the strategy outlined by Gunawan et al. (2018), where cosine similarity is utilized to quantify the degree of similarity between the news titles and the BERT embeddings of the pre-set query. The approach efficiently measures semantic similarity by taking into account the angle between vectors, which helps identify text referring to possible greenwashing activities.

Both cosine similarity and BERT aim to capture similarities between text documents or sentences, but they do so in diverse ways. Cosine similarity is a mathematical measure for measuring the similarity of texts based on their content representations, and BERT is a deep learning model for comprehending the semantic meaning of text by capturing contextual relationships. While cosine similarity gives a more straightforward and computationally efficient method of measuring document similarity, BERT can offer more sophisticated semantic understanding. Both strategies can be useful, depending on the research and the resources that are available - but their combination was used in this research given that they together can leverage the strengths of both methods. The combination of these two methods provided a more precise and contextually relevant way to compare text than using each method separately.

CEU eTD Collection

4. Chapter: Model Architecture

4.1 On the lack of greenwashing detecting tools

In 2014, Grauler and Teutenberg's survey examined consumer greenwashing detection skills (Grauler and Teutenberg 2014). Participants saw greenwashing statements alongside genuine CSR claims. In fact, the actual scores were not the key takeaway, as the survey's social media reach was more important. The goal of the creators was to raise awareness of greenwashing tactics and empower consumers to detect them from the example CSR claims. However, those already interested in sustainability skewed the survey as they were more willing to fill it out, limiting the effectiveness of it as a general awareness tool. Believing in technology instead of awareness raising, Cojoianu et al. (2020) emphasized the growing discrepancy between scientific understanding and misleading green claims which stem from the ever-growing number of fake news and misinformation on the internet. The authors pointed out climate change deniers being featured in more online articles than climate scientists, even though there was already a consensus between scientists on climate change at that time. They proposed a theoretical model in the research for a later, AI-driven framework for detecting greenwashing in this research, but their conclusion was that no effective tool could detect greenwashing, especially not across companies in different industries. They were mostly challenged by mediocre quality information and the lack of universal environmental performance indicators. These are still ongoing problems in 2024.

Three years later AI has become more widespread and more advanced, and Audun (2023) delved into the use of machine learning to automatically detect corporate greenwashing. He explored the process of collecting data from organizations using a custom web scraper. The result of this research showed that AI models, especially a "two-shot" GPT-3.5 prompt could identify greenwashing. However, fact-checking the green claims due to difficult language usage was

CEU eTD Collection

challenging, and even more importantly, the study showed that the only feasible way of obtaining relevant data currently is to scrape it from the company's website. This, in most cases, leads to the download of random information that is difficult to filter out and skews the dataset. Audun found that sophisticated tools and real-time monitoring were required to improve the suggested technique's precision. While EDHEC researchers attempted to identify greenwashing using indicators employed in climate investment strategies (Amenc, Goltz, Liu 2021), and Lueg and Lueg (2020) delved into the risks associated with sustainability reporting potentially amplifying greenwashing, their aim to develop a reliable predictive model for detecting such practices remained unsuccessful. *The exhaustive literature review revealed that current greenwashing detection methods are in the experimental phase. As of the time of this writing, no fully operational tool or model exists that adequately could detect greenwashing practices across all industries. The capability of autonomous determination is also notably absent.* A long list of failed tools could be compiled, but doing so would add little value to the discourse (the focus should be on advancing towards effective and operational solutions).

There are several smaller organizations and institutions that aim to inform the public on greenwashing via pooling together a vast amount of ESG reporting related information. However, as Lokuwaduge et al. write in their 2022 article, it has been increasingly challenging to create a comprehensive and timely solution for the credibility of ESG reporting because of its complexity, and still somewhat open interpretation (Lokuwaduge et al. 2022). Companies also use varying methodologies or even definitions, which makes the comparison between them difficult. Moreover, firms are currently permitted to cherry-pick positive information and only report that, as no enforceable regulation requires them to do otherwise (European Securities and Markets Authority 2023). While incompleteness and vagueness are red flags in ESG data and can raise

CEU eTD Collection

suspicion, additional data is required to be able to decide if the company engages in greenwashing activities. There are two problems with this outside the fact that third parties are rarely involved in ESG reporting as a controlling entity - reporting reflects historical data, and everyday consumers may not understand this specialized pool of information. Therefore, it must be concluded that using ESG data for greenwashing detection only is currently not possible.

Despite advancements in computer science and continuous attempts to develop effective models for greenwashing detection, there is still an absence of operational tools capable of determining whether a company engages in greenwashing. This of course means a gap and a significant challenge in the practical implementation and wide-spread use of detection techniques. Organizations and start-ups that were checked during the literature review usually operate as a "greenwashing data" provider and mostly use web scraping as a method of obtaining data. This, as already discussed, poses another challenge in obtaining accurate and relevant information. The proportion of random data and the difficulty of distinguishing valid and irrelevant information can distort datasets, which reduces the overall efficiency and effectiveness of techniques.

4.2 The model architecture

In the past few years, due to advancements in the field, researchers have been actively testing if different uses of AI can detect greenwashing claims from reports, articles, website content, and other publicly available information. One frequently used method for trying to get closer to greenwashing detection is the bag of words approach, first coined by Zellig Harris in 1954 (Zellig 1954). The bag-of-words model treats a text like a bag of individual words, ignoring order but keeping track of how many times each word appears (Quader, Ameen, and Ahmed 2019). This allows for turning the text into a numerical representation for use in machine learning tasks. It is a

CEU eTD Collection

conceptually and technically simple model but nonetheless has performed better than alternatives. However, while it is strong in converting text into data, it is prone to errors when working with multi-layer topics like climate change (Varini et al. 2020).

Newer models elevated to another level by natural language processing are the closest we can get to greenwashing detection as of today. An example of this is ClimateBERT, an algorithm created to assess company disclosures of climate-related risks based on TCFD. It uses a deep neural language model fine-tuned based on the language model BERT - the same one is used for the analysis of this thesis. Moreover, it also uses a two-step approach, where a machine learning model first predicts the category for each sentence within a paragraph, and then a second model predicts the category of the paragraph using the aggregated output of the first model. In the 2021 research of Bingler et al., ClimateBERT's accuracy rates were the following based on 818 TCFD reports:

- Governance: 0.94
- Strategy: 0.92
- Risk Management: 0.78
- Metrics & Targets: 0.76
- General Language: 0.71
- Overall Accuracy: 0.84

Furthermore, ClimateBERT also excelled in identifying climate-related paragraphs, as it detected climate relevance in 637 out of 653 cases, so it rarely missed climate-related text. *While there is no tool or application that can detect greenwashing at the time of writing, AI powered natural*

CEU eTD Collection

language processing models are believed to be the most effective and projected to reach complete early prediction first (Moodaley and Telukdarie 2023).

Starting with a clear explanation of the model architecture helps grasp the methodology described in the next section. Several components working together achieve greenwashing detection in this thesis. *Traditional keyword search* returns news articles related to the selected companies (Dow Jones Industrial Average companies), that contain investigations, fraud, or other news which potentially mean greenwashing. A pre-trained deep learning model, *BERT (Bidirectional Encoder Representations from Transformers)* analyzes the retrieved news article titles and generates semantic embeddings. It focuses on understanding the context and relationships between words to capture the semantic meaning of the titles. *Sentence similarity with cosine similarity* compares the BERT-generated embeddings (numerical representations) of news article titles with our predefined, literature-based criterion related to greenwashing (and a more general, dictionary definition as a baseline). Cosine similarity calculates the angle between these embeddings to determine how semantically similar the titles are to known greenwashing examples. Based on the cosine similarity score, the *"scoring model"* of the thesis assigns a "greenwashing probability score" to each company. Higher scores indicate a greater likelihood of the company engaging in greenwashing practices. Final scores can be seen in Table 5 of the Appendix.

CEU eTD Collection

Moreover, own creation of Figure 3 below demonstrates the layers of the model architecture.

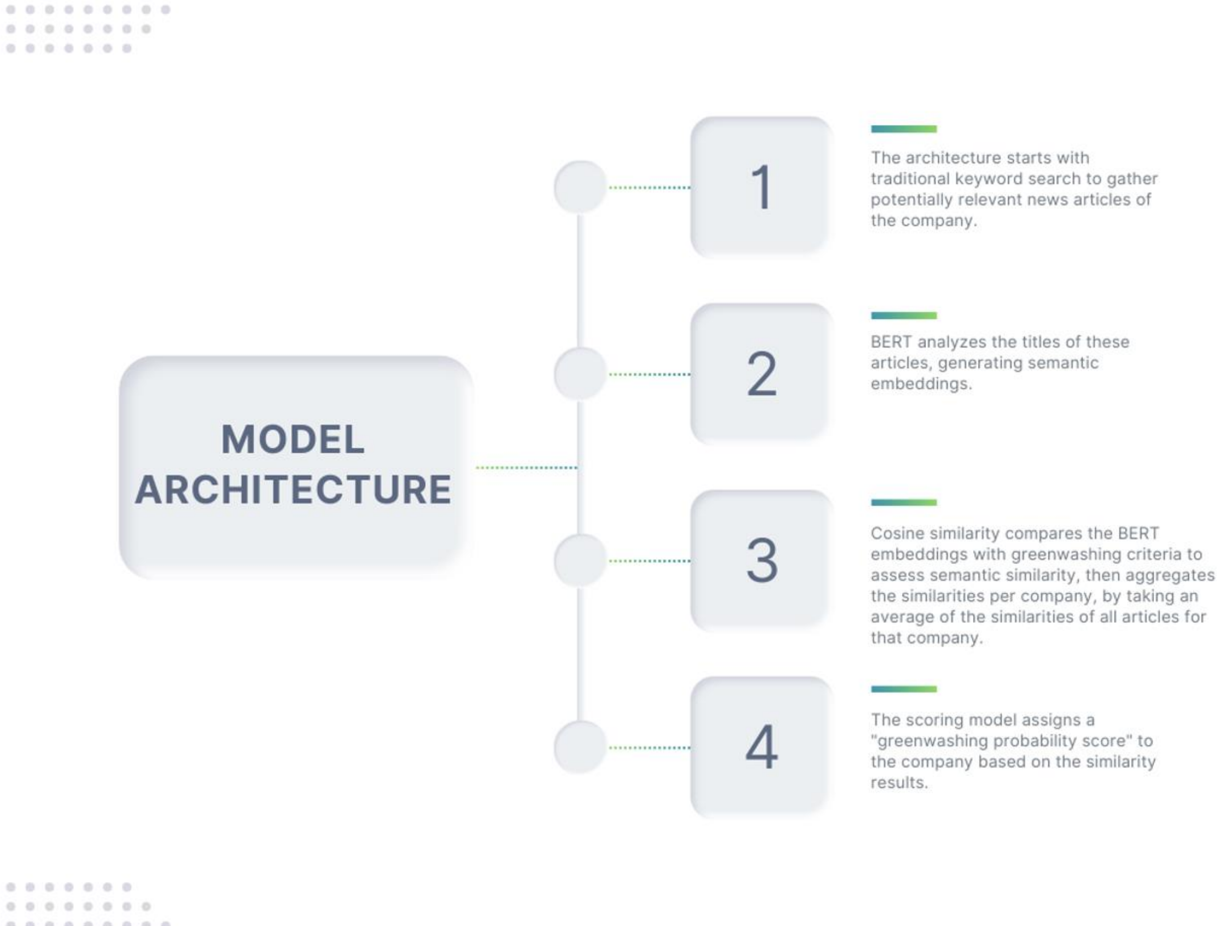

Figure 3: The model architecture of the research, showing how components get to the "greenwashing probability score" (the figure and the model architecture are the author's own creations)

This is a multi-component approach which allows for a comprehensive assessment of potential greenwashing practices. For better understanding, the code that is the engine of the model needs to be demonstrated.

CEU eTD Collection

4.3 *A look inside the code*

Although detailing the specifics of programming might distract from the research’s core purpose, sharing the code allows other researchers to understand the exact implementation of the methodology. The full code is uploaded to Github, and instead of including the entire script in the body of the thesis, a high-level overview of the code's logic and functionalities is provided. These involve the main steps of processing news articles, utilizing BERT, and generating the greenwashing probability score. The Figure 4 was created to visualize the otherwise non-transparent, hidden process that the computer does:

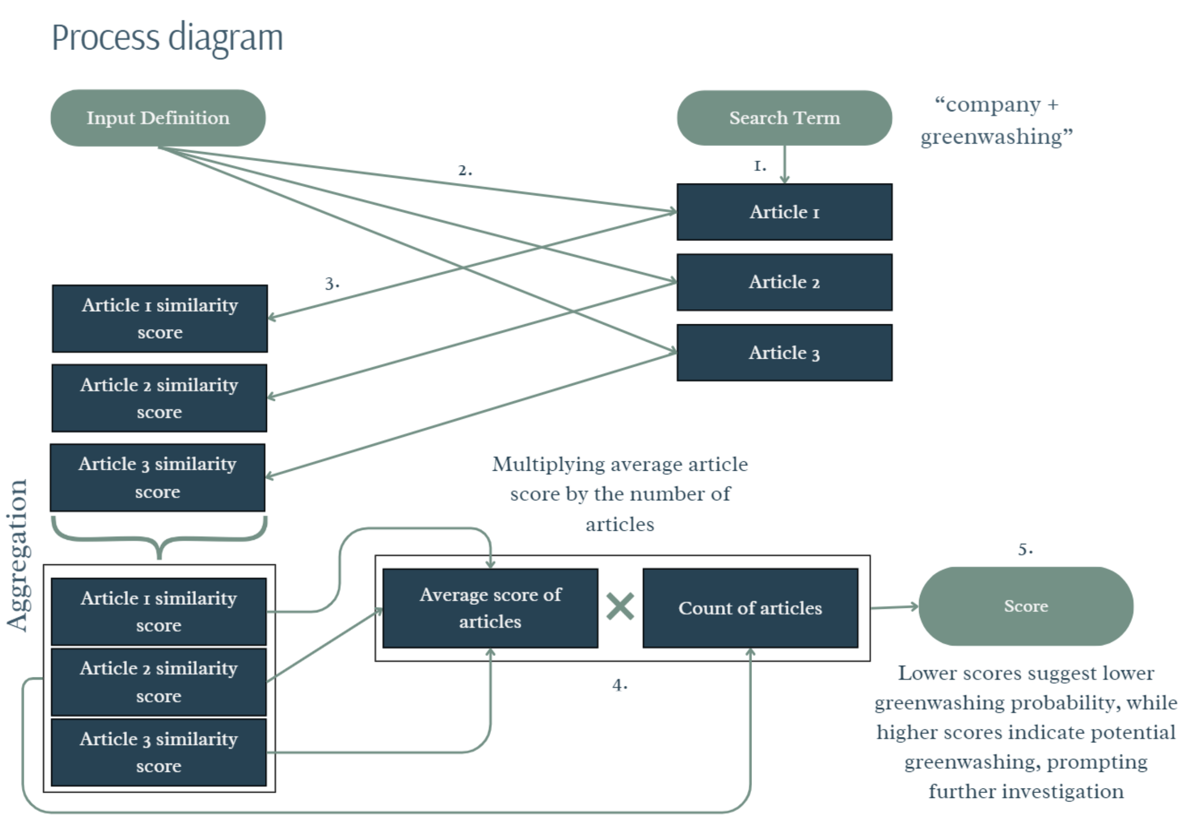

Figure 4: Process diagram showing the workings of the code (authors own figure)

CEU eTD Collection

The Python code starts by gathering potentially relevant news articles using traditional keyword search. This involves defining a keyword related to the companies that the research is investigating and potentially greenwashing practices. The keyword was the name of the firm appended by greenwashing (so “XY Company greenwashing”), as it is consistent with the information-seeking tendencies of actual consumers who would investigate a company's possible involvement in these kinds of activities. The code uses this keyword to search Google News. Once articles are retrieved, web scraping usually involves some pre-processing steps on the collected news titles, but Google News provides data that is already cleansed and in the right format.

After we have the news article titles, they are fed into the BERT, and each title goes through a two-step process. As part of transformer network processing, BERT analyzes the word order and relationships between words within the title (Zhang and Shafiq 2024). This is to capture the context and semantic meaning of the title. Then, the Siamese network structure converts the internal representation generated by BERT into a fixed-size numerical vector, which is known as a sentence embedding (Yasar 2024). This embedding essentially encodes the title's meaning in a way that can be compared to other embeddings of the sentence. The code also has predefined criteria for the four types of greenwashing practices identified based on an extensive literature review (deception and misinformation, misleading communication, selective disclosure, and decoupling).

These criteria are encoded within the script. The code utilizes cosine similarity to compare the sentence embedding generated by BERT for each news article title with the embeddings representing each of the four greenwashing criteria. A higher cosine similarity score between a title's embedding and a greenwashing criterion embedding suggests a greater semantic similarity. In other words, if the score is higher, the title likely uses language that is used in known greenwashing examples.

CEU eTD Collection

The code calculates a "greenwashing probability score" for each company by averaging the cosine similarity scores of the news article titles with the greenwashing criteria and multiplying it by the number of articles. Companies with titles that have higher similarity scores receive higher greenwashing probability scores, indicating a higher likelihood of greenwashing. For the code to function, it was necessary to import certain libraries, which are packages used for specific tasks like scraping websites to extract data and then to manipulate it. In this case, the code uses Selenium and BERT for scraping. Various Python libraries were imported as there are certain steps in the process that cannot be handled by the basic Python environment. These libraries are on Figure 5, taken from the code.

```
import time
from datetime import datetime
import pandas as pd
import tqdm as tqdm
from urllib.parse import urlparse
from selenium import webdriver
from selenium.webdriver.common.by import By
from selenium.webdriver.support import expected_conditions as EC
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.chrome.service import Service
from chromedriver_py import binary_path
from selenium.webdriver.common.keys import Keys

import random
from concurrent.futures import ThreadPoolExecutor, as_completed
```

Figure 5: The libraries used in the code

The code utilizes the "time" module and the time.sleep() function to delay execution and make sure the Google News website loads before the program interacts with the site. The datetime.now() function fetches the current date and time. Each company has its own file containing information about its articles), and the function’s main purpose is to label the Excel files saved after each search. The addition of timestamps improves the organization and

CEU eTD Collection

traceability of the output data, which is particularly crucial for subsequent analysis. While not mandatory for data collection, two libraries simplify interactions with Google News in the model. The "Pandas" data manipulation library organizes data in tabular form (rows and columns) and structures it into data frames, which are then saved as Excel files. "TQDM" (not explicitly used) tracks progress during data iteration and acts as a self-checking mechanism. The module "urllib.parse" deals with parsing and structuring URLs. It is used to extract the domain name from Google News URLs when performing web scraping.

"Selenium" empowers Python to interact with web browsers. This testing library is a popular choice for web scraping tasks, in the research's model, it interacts with Google News, performs searches, and retrieves the desired information. The "CromeDriver" library essentially ensures compatibility between Chrome and Selenium. The "keys" class, which is also part of Selenium, simulates pressing keyboard keys. Specifically, pressing the "Enter" key after entering a search term (just like a person would do so in real life). "SentenceTransformers" is the library for the previously introduced BERT model. The "util" module from SentenceTransformer provides additional utility functions, specifically calculating cosine similarity between the embeddings of predefined search terms and Google News article titles.

These libraries work together to allow the script to do a variety of tasks, from web scraping to natural language analysis. They are the tools needed for efficient and objective data collection but also minimize any biases and manual errors common with manual data collecting.

CEU eTD Collection

BERT was set up via the SentenceTransformers Python library, and to work, it needed a predefined statement which it could encode into an embedding. This question was manually set, and for the main analysis of the research, it was “This company is or was involved in a greenwashing scandal where they purposefully deceived their customers." When computing, BERT looked at Google News articles and compared the articles of those to this statement, finding similarities. Just to demonstrate this through a fictional example, if Inditex (the fashion giant owning Zara), had a greenwashing scandal about a collection that was claimed to be sustainable but in fact was not, 5BERT gave the company a high score. Let us imagine two title versions: *“Zara customers outrage as new collection turns out not to be made of recycled materials”* and *“Zara tricked customers into buying “recycled clothes,” again”.* In this case, BERT is well-trained enough to know that contextually and semantically, the second version is indeed more like the predefined statement, therefore, it would give it a higher similarity score.

The articles were web scraped, which means the automated process of obtaining data from a website. With less and less people reading print media (Chen 2023), greenwashing scandals are most of the time covered by online news media companies. In this section of the research, it is important to discuss the web scraping process itself, as this is a highly customizable process but there is no golden standard that one should follow. For this research, first, the target website was defined as Google News. For the main analysis, this was the *link* of the search results for the company name appended by greenwashing, as shown below on Figure 6.

CEU eTD Collection

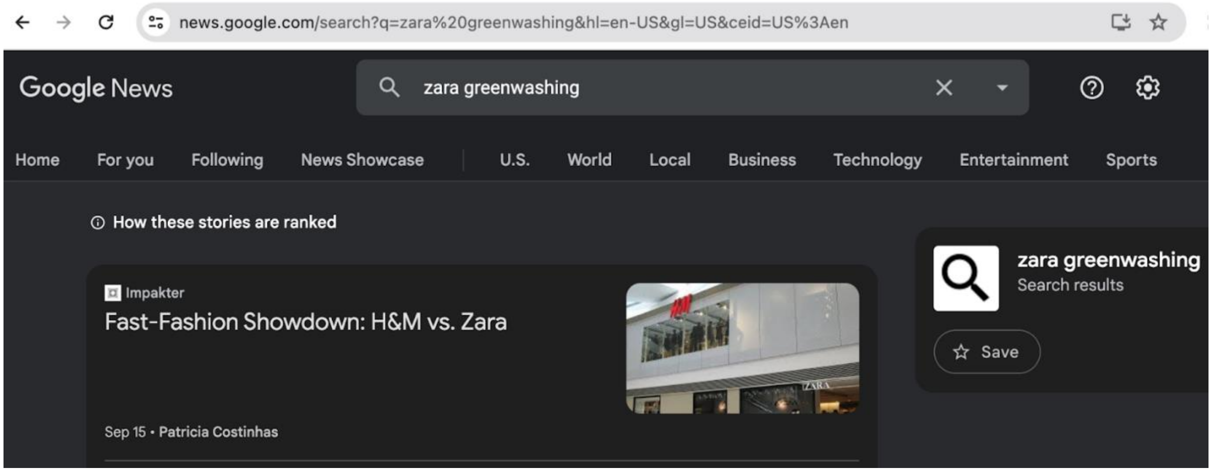

Figure 6: Google News search example

The next step of the program was to open the article by clicking on it, then find and copy the article title and the date of publishing. Then, it had to paste this information into an Excel file. The program repeated this process for each article of the search results, compiling the "company's Excel file" containing all articles that can possibly be connected to the given company's greenwashing activities.

Altogether, 30 companies' 1131 articles were found and analyzed in the research. Figure 7 and figure 8 visualize the sum of articles by company and by industry. The first chart indicates that companies like Coca-Cola, Apple, and Amazon have the highest number of articles, suggesting they are under scrutiny for greenwashing practices. The second chart shows that information technology (19.45%) and financial services (13.88%) dominate the industry distribution.

CEU eTD Collection

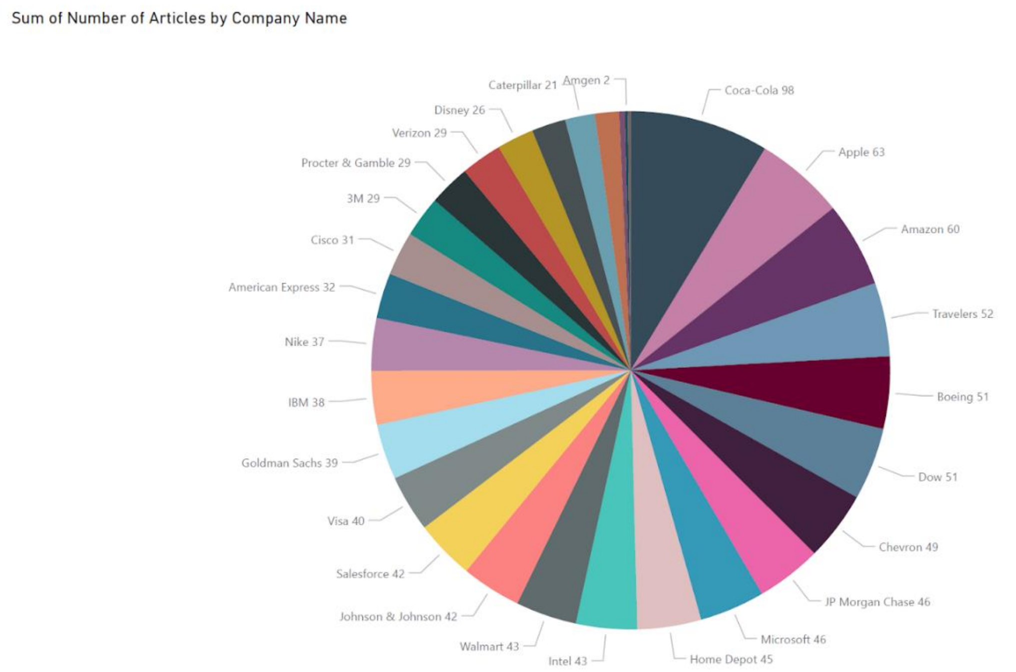

Figure 7: Pie chart showing each company and the sum of articles (author's own chart, created in Power BI)

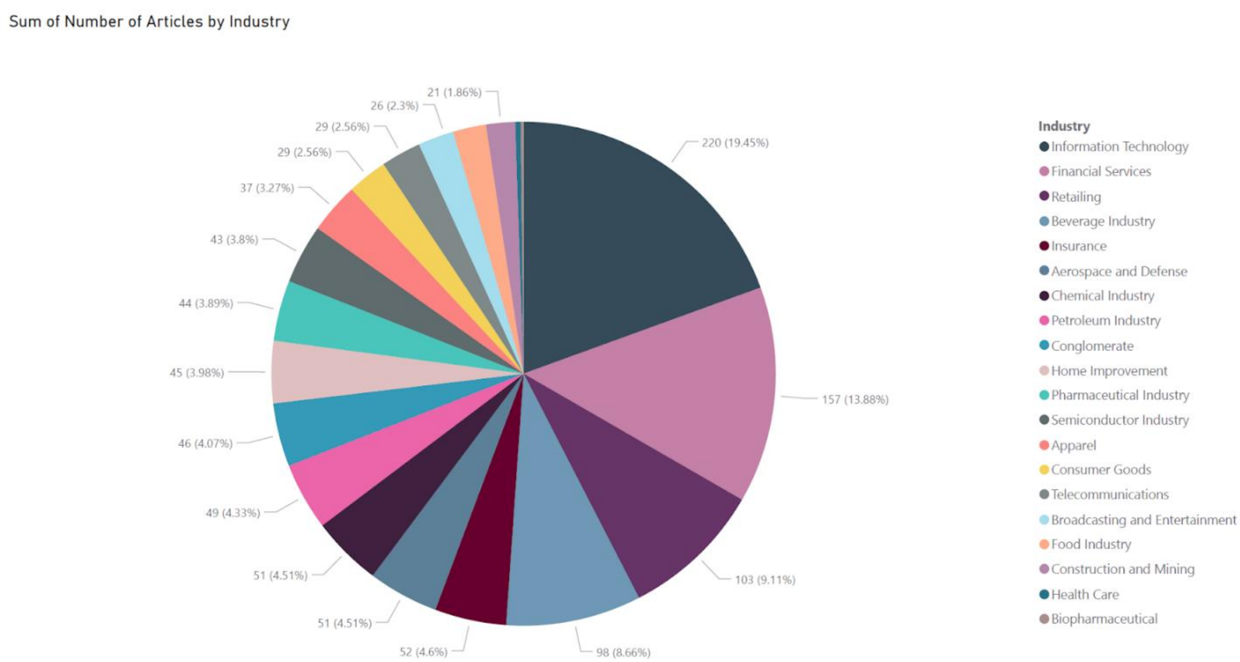

Figure 8: Pie chart showing each industry and the sum of articles (author's own chart, created in Power BI)

CEU eTD Collection

There are multiple elements from which the web scraping process comes together. For a clearer code explanation, the following table (Table 1) was created, which lists all the code elements that were used, and their descriptions:

Step	Code Element	Description
1.	driver = webdriver.Chrome(...)	Initializing Chrome browser using Selenium.
2.	driver.get(google_news_page)	Opening the Google News homepage in the browser.
3.	WebDriverWait(driver, 10).until(...)	Using WebDriverWait to wait 10 seconds for elements to become clickable before interacting with them.
4.	input_field = WebDriverWait(driver, 10).until(...)	Locating and waiting 10 seconds for the search input field to become clickable.
5.	input_field.send_keys(file_name + Keys.ENTER)	Entering the search term and simulating pressing 'Enter' in the search bar of Google News.
6.	title_elements = driver.find_elements(By.CLASS_NAME, 'JtKRv')	Locating elements with class 'JtKRv', which represent the article titles on the Google News search results page.
7.	date_elements = driver.find_elements(By.CLASS_NAME, 'hvbAAd')	Locating elements with class 'hvbAAd', which represent article publication dates.

CEU eTD Collection

8.	link = title_element.get_attribute('href')	Extracting the URL (link) of the article from the title element.
9.	title_element.text	Getting the text content of the title element, representing the article title.
10.	date_element.get_attribute('datetime')	Extracting the publication date of the article in a standardized format.
11.	pd.DataFrame(articles)	Converting the collected data (titles, links, dates) into a Pandas DataFrame (articles_df).
12.	model.encode(title)	Use SentenceTransformer to encode the article title into a numerical embedding.
13.	util.pytorch_cos_sim(question_embedding, title_embedding)	Find the cosine similarity between the embedding of the predefined question and the embedding of the article title.
14.	articles_df['Similarity Score'] = similarity_scores	Add the calculated similarity scores to the DataFrame.
15.	articles_df = articles_df.sort_values(by='Similarity Score', ascending=False)	Sort the DataFrame based on similarity scores in descending order.
16.	articles_df.to_excel(file_name_final, index=False)	Save the final DataFrame to an Excel file.

Table 1: Code elements and their function

CEU eTD Collection

4.3.1 Developing a custom ranking system to weigh greenwashing better

The resulted raw similarity score might not tell the whole story. There are two reasons why ranking had to be used (which is similarity multiplied by the number of articles). As mentioned already, the model utilizes cosine similarity to assess how semantically similar news article titles are compared to predefined greenwashing definitions. Imagine that Company A has one article with a remarkably high cosine similarity score (strong indication of greenwashing), and Company B has ten articles with a moderate cosine similarity score. The two must be measured against each other because the ten "weaker" articles of Company B overall can mean a stronger greenwashing potential. By multiplying similarity by the number of articles, the ranking system prioritizes Company B for further investigation despite the lower individual similarity scores. This approach ensures that companies with a higher volume of potentially greenwashing news coverage are flagged.

The ranking also addresses potential bias. If the greenwashing criteria would rely on keywords or specific phrases, there would be a risk of companies strategically avoiding those exact terms while still engaging in greenwashing practices, as is becoming more and more common in real life. The number of available articles can help address this. Even if a company avoids specific keywords, a high number of articles discussing their environmental practices might still suggest potential greenwashing, prompting further investigation. Therefore, using ranking can provide a more nuanced picture of potential greenwashing compared to relying solely on cosine similarity scores. It helps prioritize companies for investigation based on both the strength of the greenwashing signal and the volume of potentially relevant news coverage. Final scores after raking can be seen in Table 6 of the Appendix.

CEU eTD Collection

5. Chapter: Results

This section is the single most important part of the thesis as it presents the findings of the analysis conducted using the developed NLP-enhanced framework.
A spreadsheet in Excel was created by running the code. It contained weighted scores for "greenwashing" for each company by each complied definition. The PERCENTILE.EXC function was utilized to determine the 15th percentile as a percentage. Subsequently, online research was conducted for companies that exceeded this threshold (n=12) to validate the existence of greenwashing cases. Out of the twelve companies, eleven were found to have engaged in greenwashing, but this level of precision does not directly mean a 91.67% success rate for the proposed framework. To get the real results, a range of widely used machine learning metrics were utilized, including a confusion matrix. This matrix offered valuable insights into the model's performance by classifying companies into four key groups (Tiwari 2022). True positives (TP) are the companies that were predicted to be above the threshold and had confirmed greenwashing. False positives (FP) are the companies that were predicted to be above the threshold but had no confirmed greenwashing. True negatives (TN) are the companies that were predicted below the threshold and had no confirmed greenwashing. Lastly, false negatives (FN) are the companies that were predicted below the threshold but had confirmed greenwashing. These four confusion matrix elements could be counted with the help of the code's output table, shown in Table 7 of the Appendix.

Using the confusion matrix, performance metrics were calculated to statistically evaluate the model. The model's accuracy reflects how correct it is overall, showing the ratio of correctly predicted instances to the total number of companies. Precision is the proportion of identifications that were truly greenwashing. In our case, precision indicates the model's ability to accurately

CEU eTD Collection

predict greenwashing cases. Recall is the ratio of true positives to all actual greenwashing cases, measuring the capacity of the model to capture the “entirety” of existing greenwashing practices. The F1 Score computes the harmonic average of precision and recall, providing a balanced assessment of the model's effectiveness.

These metrics are all presented and discussed in detail below, along with a critical analysis of the model's performance and limitations. To set the context, however, a summary of the main findings is listed already here. First of all, the model is *statistically* significant, meaning it is not just identifying greenwashing by random chance. The proposed NLP-enhanced machine learning model can effectively identify greenwashing with *high precision* (88-93% depending on the definition). This means that the model is effective in correctly identifying predicted greenwashing cases. The model's performance, however, *varies* depending on the specific greenwashing approach analyzed. *It performs better at identifying certain greenwashing tactics* (like selective disclosure) than others (like deception and misinformation). The model also *misses* a substantial portion of actual greenwashing cases (recall is only 55-75% depending on the definition). This means there is room for improvement in capturing all greenwashing cases. Finally, certain industries (such as information technology and financial services) are suspected to be more prone to greenwashing than others (like pharmaceuticals), according to the model.

CEU eTD Collection

5.1 Confusion matrix analysis

The model's predictions regarding each company were classified based on two criteria: (A) exceeding a predetermined threshold and (B) confirmation of past or present greenwashing practices. Keeping this classification in mind, the number of companies for each of the four categories could be counted.

The model successfully identified 11 companies that are actually engaged in greenwashing. These include notable Dow Jones companies such as Coca-Cola, Apple, Amazon, Microsoft, Chevron, Boeing, Walmart, Dow, JP Morgan Chase, Nike, and Johnson & Johnson. However, the model missed 9 companies that are also engaging in greenwashing practices; these are Visa, Procter & Gamble, Home Depot, IBM, Goldman Sachs, McDonalds, Disney, Verizon, and Caterpillar. Additionally, there was one instance where the model incorrectly flagged a company, Travelers Companies Inc., as a greenwasher when it was not. On a positive note, the model accurately identified 9 companies as not engaging in greenwashing, which are Salesforce, American Express, Intel, Cisco, 3M, Honeywell, UnitedHealth Group, Amgen, and Merck & Co. Inc. The summary of numbers can be seen on the below confusion matrix (please note that this confusion matrix is only for the “general greenwashing definition” category, but the results of the other approaches will also be detailed below):

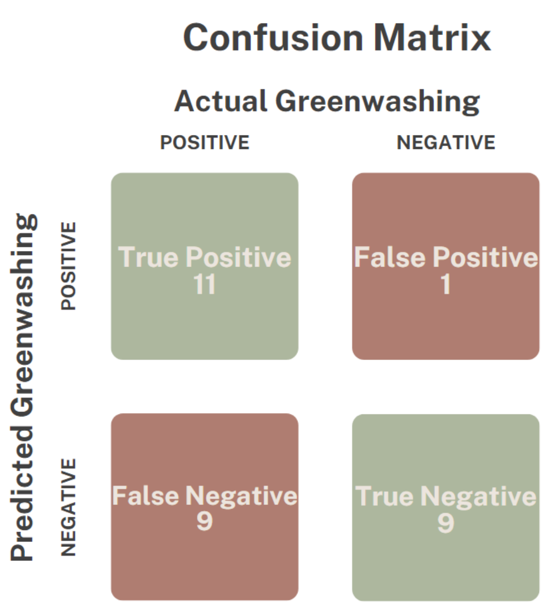

Figure 9: Confusion matrix for the "general definition" (author's own matrix)

CEU eTD Collection

From the confusion matrix, not only binary classifications can be deducted, but key performance metrics and insights too, which evaluate the effectiveness of the model. The following are the most significant metrics in terms of model performance: accuracy, precision, recall, and F1 Score.

5.2 Performance metrics

As can be seen from the figure below, creating the thesis's confusion matrix is useful since it allows for the deduction of different performance indicators from its four categories.

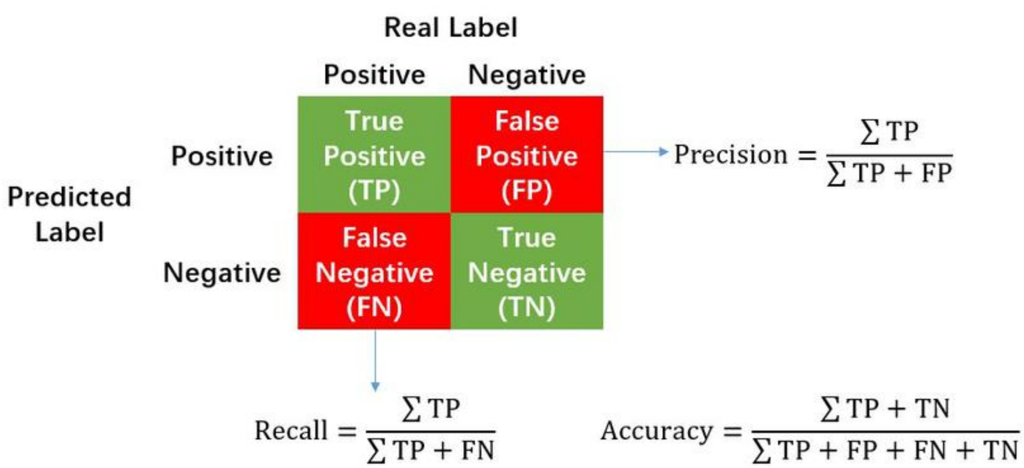

Figure 10: Calculating main metrics from the Confusion Matrix, adapted from MA et al. (2019), 'Analyzing the Leading Causes of Traffic Fatalities Using XGBoost and Grid-Based Analysis: A City Management Perspective'

The model's *accuracy* refers to its overall correctness, which is the proportion of correctly predicted instances to the total instances. It is important to note that true negatives are as important as true positives. The accuracy result is:

$$\frac{TP + TN}{TP + TN + FP + FN} = \frac{11 + 9}{11 + 9 + 1 + 9} = \frac{20}{30} \approx 0.666$$

CEU eTD Collection

This means that the model correctly identifies 66% of the cases. While this is reasonable accuracy, for a balanced evaluation, this metric alone might not be enough.

The proportion of positive identifications that were correct is known as *precision*. It is also the ratio of true positives to the total predicted positives. It indicates how many of the predicted greenwashing cases were actual greenwashing. The result is:

$$\frac{TP}{TP + FP} = \frac{11}{11 + 1} = \frac{11}{12} \approx 0.917$$

This means that the model is highly precise, when it predicts greenwashing, it is correct 91.7% of the time.

Recall is the ratio of true positives to the total actual positives. It measures the model's ability to correctly identify all actual greenwashing cases.

$$\frac{TP}{TP + FN} = \frac{11}{11 + 9} = \frac{11}{20} \approx 0.55$$

The recall is moderate, the model misses a substantial portion of actual greenwashing cases, 45%. There is room for improvement.

The F1 Score combines precision and recall into one balanced measure, also demonstrating the trade-off between the two (Sharma 2023).

CEU eTD Collection

$$F1\ Score = 2\ \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0.917 \times 0.55}{0.917 + 0.55} \approx 0.687$$

An F1 score of 68.7% indicates that the model performs reasonably well, but there is room for improvement, particularly in increasing recall to capture more actual greenwashing cases.

However, the purpose of the thesis is not necessarily to identify more greenwashing cases (i.e., improve recall), but rather to confirm greenwashing when suspected. While the model might not be able to predict all greenwashing cases, it correctly predicts greenwashing 91.7% of the time.

5.3 Hypothesis testing with Chi-square

The Chi-square test is used to evaluate independence and association (between model predictions and confirmed greenwashing cases). In other words, to compare expected results with observed ones. It is used for testing hypotheses, as it can show discrepancies between observed and expected results (Hayes 2024).

Category	Observed	Expected
Confirmed greenwashing	11	9
No confirmed greenwashing	1	9

Table 2: Observed and expected greenwashing cases

Calculations for the two categories are as follows:

Confirmed greenwashing: $\frac{(O-E)^2}{E} = \frac{(11-9)^2}{9} \approx 0.444$

CEU eTD Collection

No confirmed greenwashing: $\frac{(O-E)^2}{E} = \frac{(1-9)^2}{9} \approx 7.111$

The sum of these two is the test statistics: $x^2 = 0.444 + 7.111 = 7.555$

From here, the CHISQ.TEST Excel function was used, as it requires minimal input, reducing the risk of errors in manual calculations. Using this Excel formula, the p-value of the model resulted in 0.005982535, rounded to 0.006 or 0.6%. In our situation, a confidence interval of 95% was established, indicating a significance level of 5%. This percentage, 5%, is frequently used in such assessments (Berkson 1942).

The interpretation of the Chi-square test is as follows. According to the calculated p-value, there is a 0.6% chance of observing a test statistic as extreme as 7.555 under the null hypothesis. In other words, if the research's model did not significantly detect greenwashing (this is the null hypothesis), the probability of getting a test result as large as 7.555 due to random chance is only 0.6%. This leads us to the following finding:

The p-value is less than the significance level (0.6% < 5%), so the null-hypothesis of the research can be rejected.

In other words, if the model did not actually detect greenwashing effectively, the probability of seeing such a strong alignment between the model's predictions and the actual greenwashing cases purely by chance would be extremely low, 0.6%. Because this probability is very low, it can be concluded that the observed alignment is *not* due to random chance.

CEU eTD Collection

Figure 11 summarizes p-score ranges and how strong the evidence is against the null hypothesis. Based on this practice, the 0.6% probability is significantly lower than the standard 5% threshold, providing strong evidence against the null hypothesis. Consequently, we can reject the null hypothesis and conclude that the model effectively identifies greenwashing. This indicates that the differences between the observed and expected frequencies are not due to random chance and that our model's predictions are meaningfully aligned with the actual greenwashing status.

Values of p	Inference
$p > 0.10$	No evidence against the null hypothesis.
$0.05 < p < 0.10$	Weak evidence against the null hypothesis
$0.01 < p < 0.05$	Moderate evidence against the null hypothesis
$0.05 < p < 0.001$	Good evidence against null hypothesis.
$0.001 < p < 0.01$	Strong evidence against the null hypothesis
$p < 0.001$	Very strong evidence against the null hypothesis

Figure 11: Different p-value inferences, adapted from Singh (2013), 'P Value, Statistical Significance and Clinical Significance'

Coming back to performance metrics, accuracy of 66.6% provides a general overview, precision 91.7% indicates that the model accurately identifies most predicted greenwashing cases. The F1 score of 68.7% offers a balanced view. However, recall of 55% shows that there is room for improvement in capturing all actual greenwashing cases.

The Chi-square test confirms a statistically significant association (p-value = 0.006) between the predictions of the model and confirmed greenwashing cases. This allows for the rejection of the null hypothesis, indicating that the proposed framework effectively detects articles on actual greenwashing practices among the studied companies. Overall, the model demonstrates a

CEU eTD Collection

promising performance in identifying greenwashing with high precision and statistically significant results. Nonetheless, as it is explained later in the research, the proposed model does come with limitations.

5.4 Visualizing the results

As already discussed in the methodology section, the thesis fills a gap in existing literature: the greenwashing definitions used to conduct research are in most cases random, generalized, or focused on one aspect. In this research, different greenwashing definitions focus on all areas that are prevalent in research on greenwashing. As a reminder, the four identified definitions besides the broadest, general one are deception and misinformation, misleading communication, selective disclosure, and decoupling. All four focus on different aspects of greenwashing and they differ in the emphasis they place on the deceptive intent and the methods used. One key result to notice is that, as Table 3 shows, the p-values varies across the different methodologies in the analysis:

Greenwashing approach	TP	TN	FP	FN	P-value
General definition	11	9	1	9	0,005982535
Deception and misinformatior	11	9	1	9	0,005982535
Misleading communication	14	9	1	6	0,000024827
Selective disclosure	15	8	2	5	0,000000743
Decoupling	14	9	1	6	0,000024827

Table 3: Variance of p-values between different greenwashing approaches (author's own calculations)

As the second finding of the research, the variance in p-values between different greenwashing definitions suggest that the model performs better at identifying certain greenwashing tactics than others, based on the features it analyzes. This aligns with the nuanced

CEU eTD Collection

nature of greenwashing, given that companies employ various deceptive communication strategies. The statistically significant (p-value < 0.05) association between the model's predictions and actual greenwashing classifications for all approaches 1) supports the existence of these classifications and 2) demonstrates the model's potential to detect them.

Misleading communication and decoupling have lower p-values (0.000024827) compared to the general definition (0.005982535 or 0.6%) and deception and misinformation. The lower p-values suggest that the model is more effective at recognizing specific greenwashing tactics based on the features it analyzes. This does not necessarily mean that these tactics are "worse" than others, rather, the model is better equipped to identify them due to its architecture and the data on which it was used. For example, sustainability campaigns or symbolic actions might be flagged by features that identify vague claims or a disconnect between messaging and actions.

Selective disclosure has the lowest p-value (0.000000743), indicating a very strong association between model predictions for selective disclosure and actual greenwashing classifications. This suggests that the model is particularly good at identifying cases where companies selectively disclose positive environmental information. There are theoretical reasons why selective disclosure is better recognized by the model, and why it resulted in the lowest p-value. Selective disclosure in most cases involves using specific language patterns like "committed to sustainability," or highlights of participation in sustainability programs without mentioning actual impact. These are types of patterns that the model may easily capture. The model may also recognize selective disclosure better because corporations are more likely to utilize this approach because it is subtler and more difficult to definitively debunk (prove scientifically).

CEU eTD Collection

Extending on what was said about the performance metrics of the general greenwashing definition of the research, Table 4 contains the results of all other evaluated greenwashing approaches:

Greenwashing approach	TP	TN	FP	FN	P-value	Accuracy	Precision	Recall
General definition	11	9	1	9	0,005982535	0,666667	0,916667	0,55
Deception and misinformation	11	9	1	9	0,005982535	0,666667	0,916667	0,55
Misleading communication	14	9	1	6	0,000024827	0,766667	0,933333	0,70
Selective disclosure	15	8	2	5	0,000000743	0,766667	0,882353	0,75
Decoupling	14	9	1	6	0,000024827	0,766667	0,933333	0,70

Table 4: All examined greenwashing approaches' statistical metrics

Just like the p-value, accuracy, precision, and recall resulted in varying values. Deception and misinformation had the same values as the general definition. The highest results were achieved by misleading communication and decoupling, while the lowest by selective disclosure.

Several charts have been prepared in a data visualization software called Power BI, to allow for visualizing the results and the easy deduction of potentially important information. These charts directly support the key arguments or findings of the research, so it was beneficial to include them in the main body instead of the appendix. They must be included in the main body so that the reader can quickly understand the data, conclusions, and supporting evidence.

Figure 12 visualizes the performance of the 5 greenwashing approaches. Precision is highest for decoupling and misleading communication (0.93) and lowest for selective disclosure (0.88). Accuracy is consistent for decoupling and misleading communication (0.77) but is lower for deception and misinformation, and general definition (0.67). Recall is highest for selective disclosure (0.75) and lowest for deception and misinformation, and general definition (0.55). The chart also shows that across all approaches, the model performs best in terms of precision, but

CEU eTD Collection

particularly for decoupling and misleading communication. However, it struggles with recall for deception and misinformation, and general definition.

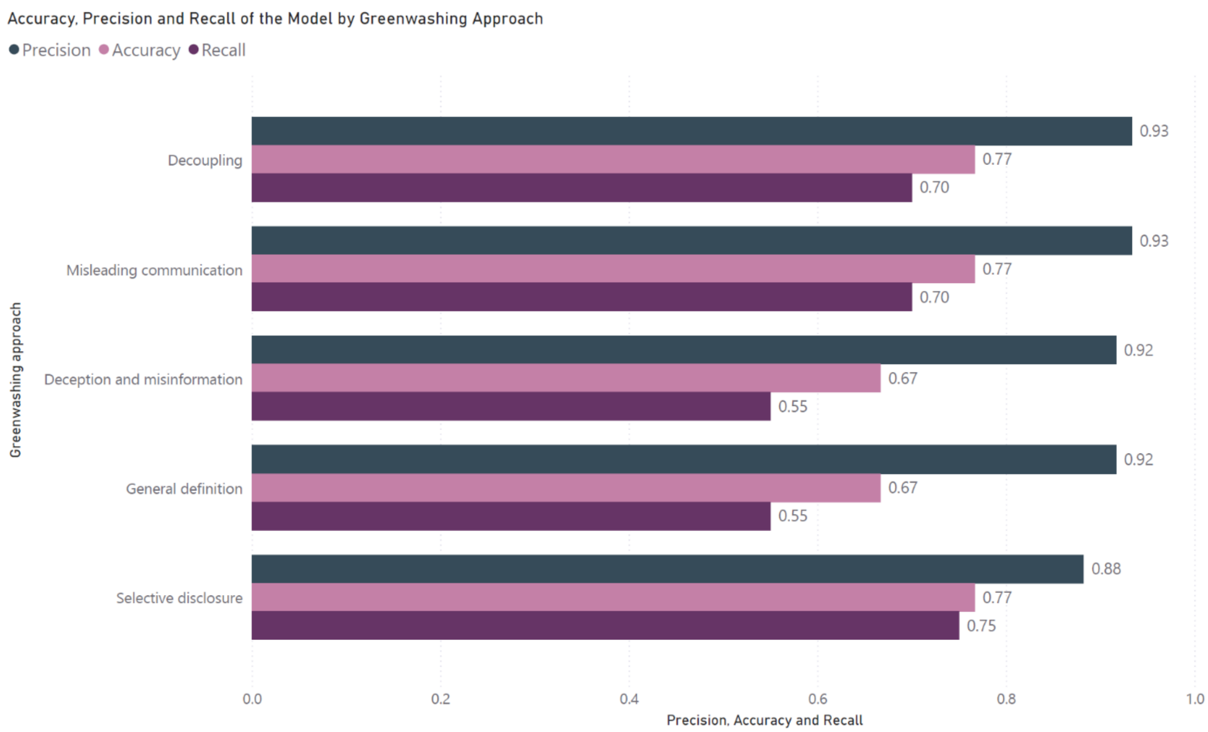

Figure 12: Bar chart of accuracy, precision, and recall by greenwashing approach (author's own chart)

Figure 13 visualizes the distribution of false negatives, false positives, true positives, and true negatives among the analyzed greenwashing approaches. This visualization provides a clearer picture of the model performance in terms of actual counts of various predicted outcomes for each

CEU eTD Collection

approach.

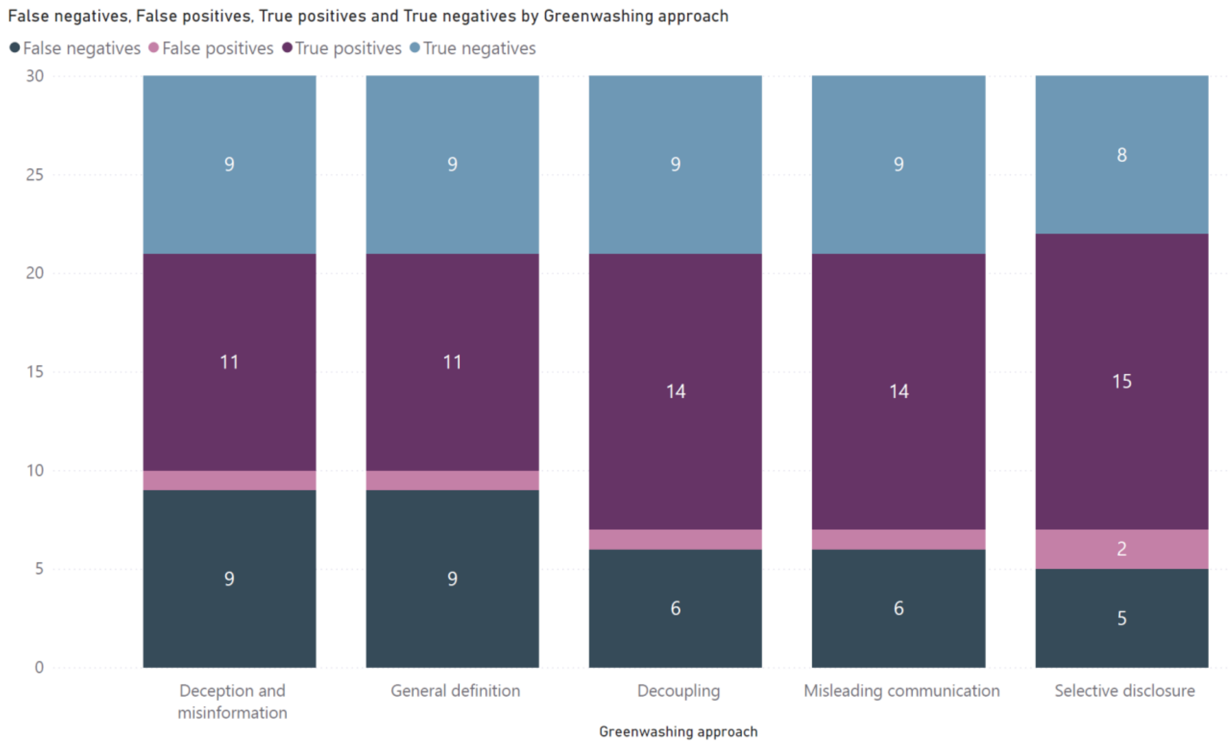

Figure 13: Clustered chart of FN, FP, TP, and TN by greenwashing approach (author's own chart)

Its interpretation is as follows. The ”deception and misinformation” and ”general definition” approaches have a higher number of false negatives, indicating that the model more often misses these instances. Selective disclosure has the highest true positives and the lowest false negatives, suggesting that the model performs best for this greenwashing approach. What is even more important is the fact that the consistency of false positives (mostly 1) and true negatives (mostly 9) across most approaches shows that the identification of true negatives is reliable, with minimal positive identifications that are incorrect.

Figure 14 shows the distribution of false negatives and false positives by greenwashing approach:

CEU eTD Collection

Figure 14: False negatives vs false positives (author's own chart)

The combination of high false negatives (35) and low false positives (6) on the blue bars suggests that the model is more conservative in identifying greenwashing instances, which leads to missed true positives (false negatives) but fewer incorrect identifications (false positives). Out of all categories, selective disclosure stands out with its balance since it has both low false negatives and low false positives. This indicates that the model has better performance in identifying this type of greenwashing correctly. Deception and misinformation and general definition have the largest false negative rates, indicating the model's difficulties in effectively capturing these, as shown by their lower recall values.

Figure 15 and 16, the two below ribbon charts both compare how different companies stack up against each other in terms of their greenwashing practices. The first one is for the thirty Dow Jones companies, but the second one introduces a new category: industries of the Dow Jones

CEU eTD Collection

companies. The width of each ribbon along the x-axis represents the intensity (score) of greenwashing practices. Wider sections therefore indicate higher scores, suggesting more prevalent greenwashing practices.

Companies and industries higher up on the chart are ranked worse, suggesting they engage in more greenwashing. The industry chart starts with information technology, financial services, and beverage industry with the widest parts of their ribbons. This indicates that these industries, on average, have higher greenwashing scores compared to others like healthcare or pharmaceutical towards the end of the chart. The color layers help identify which greenwashing tactics are most common in each industry. The tapering of the ribbons indicates that while some companies within an industry may score high on greenwashing, others do significantly better, which leads to a decrease in ribbon width as fewer companies in the sector engage in these practices.

CEU eTD Collection

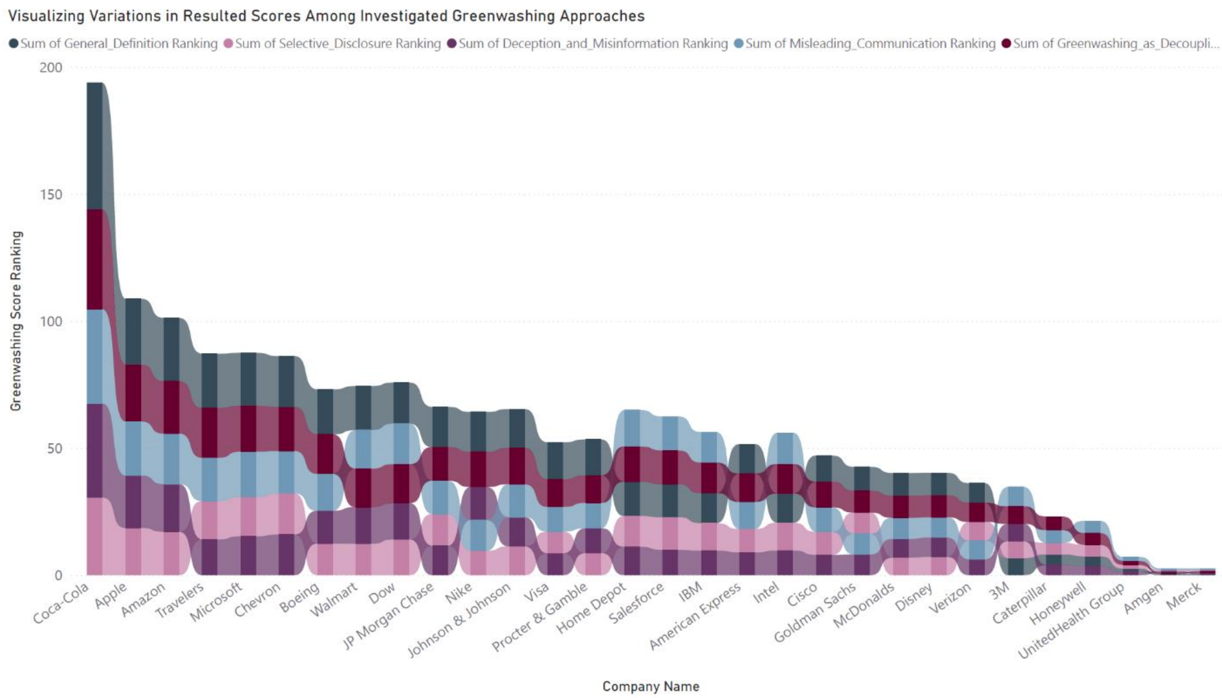

Figure 15: "Greenwashing score" variation among companies and among greenwashing approaches (author's own chart)

CEU eTD Collection

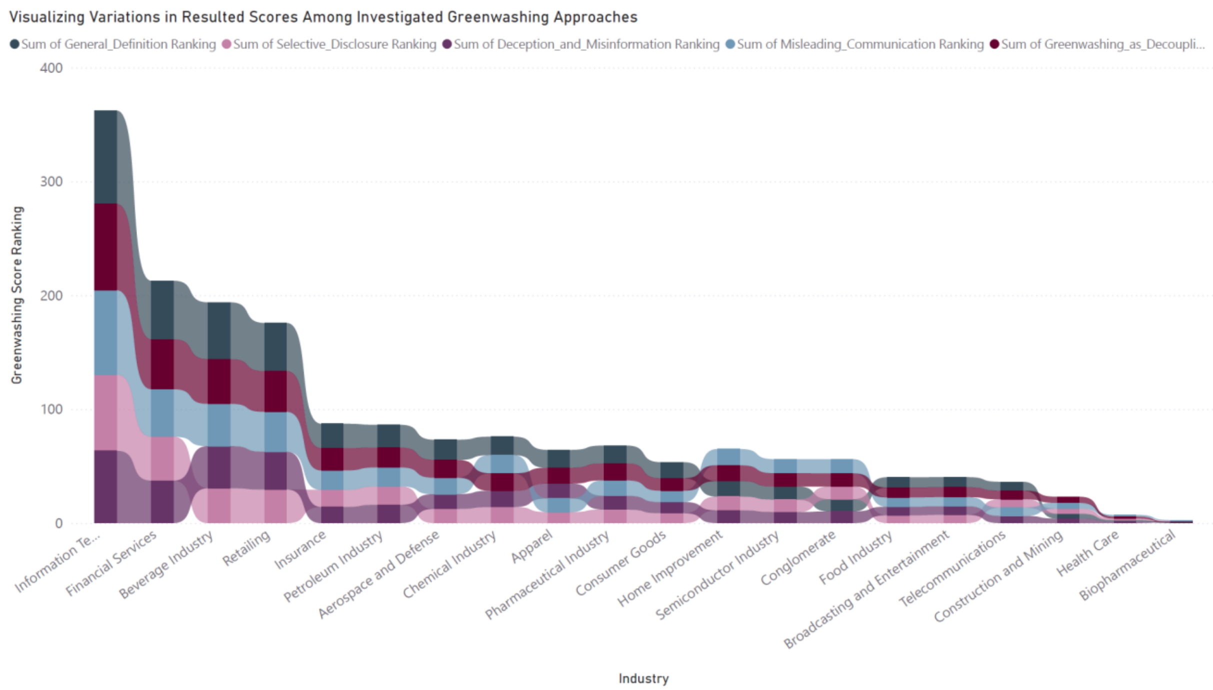

Figure 16: "Greenwashing score" variation among industries and among greenwashing approaches (author's own chart)

The last visualization, Figure 17 is a tree map that illustrates the sum of general definition scores for different industries. The size of each rectangle represents the total score calculated by the model for that industry. This provides a view where industries can easily be compared visually.

CEU eTD Collection

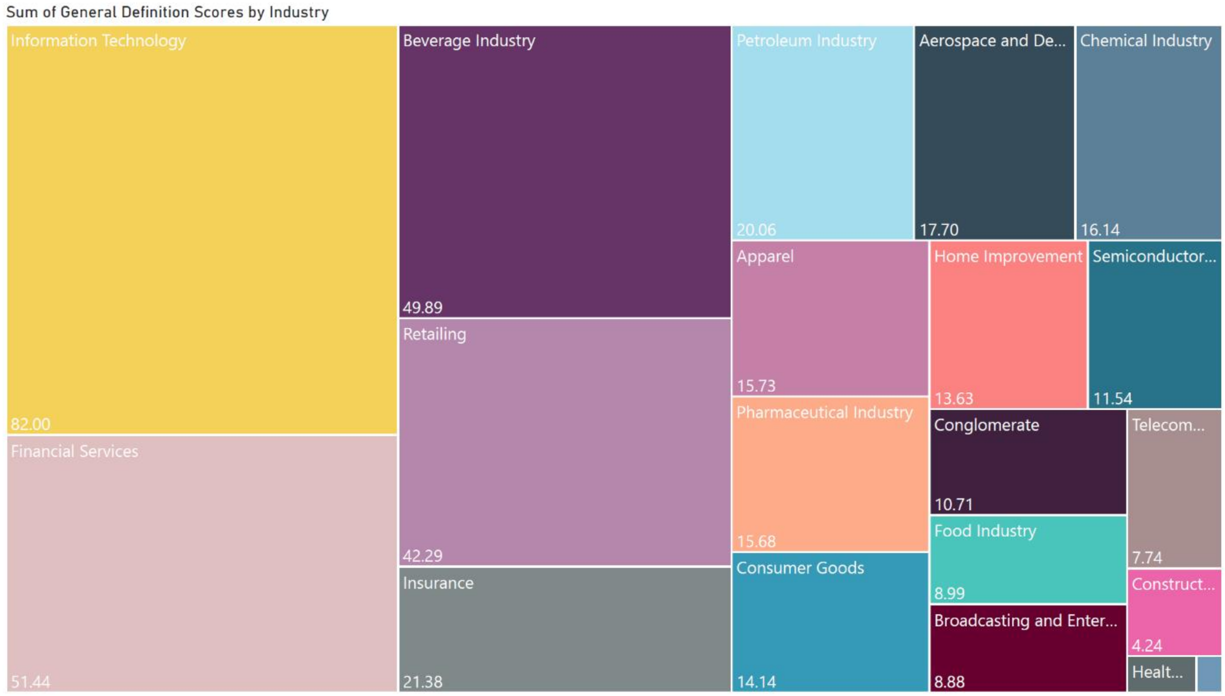

Figure 17: Greenwashing potential of industries (author's own chart)

Information technology and financial services lead the tree map, with scores of 82.00 and 51.44, respectively. These industries have the highest potential for greenwashing based on the model's scoring system. With scores of 49.89 and 42.29, the beverage and retailing industries also show significant potential for greenwashing. They are notable but less prominent compared to the top two sectors. Scores of 20.06 and 21.38 suggest moderate levels of potential greenwashing for the petroleum and insurance industries. These sectors are important to watch but do not dominate the landscape as the industries so far. The aerospace and defense, chemical, and apparel industries have lower scores ranging from 15.73 to 17.70, indicating a potential for greenwashing activities. Sectors like pharmaceutical, semiconductor, consumer goods, conglomerate, and telecommunications have even lower scores, meaning they might be less involved in greenwashing, according to the framework.

CEU eTD Collection

The results overall demonstrate that the proposed framework does identify greenwashing effectively with high precision and statistically significant results. The model's performance varies across different greenwashing definitions, but this highlights the complexity of corporate greenwashing. The visualizations that were created in Power BI support the findings, as they not only emphasize the prevalence of greenwashing in certain industries but also the model's ability to capture those nuances that cannot be seen only from an Excel sheet. Despite having these results, however, there are areas that could be improved, especially in terms of recall. In the following chapter, the thesis's findings and contributions are discussed, as well as the implications for future research, and current limitations.

6. Chapter: Discussion

6.1 Discussing results

The first component of the framework, BERT evaluates how well news article titles match the four predetermined greenwashing criteria in a semantic sense. The second component of the framework, namely the volume of articles, aims to provide the necessary context of occurrence frequency. The overall framework aims to predict those companies that may have a higher probability of greenwashing based on the product of their article numbers and respective similarity scores. This method guarantees thorough coverage aiming to expose businesses that engage in widespread greenwashing, consequently narrowing down the scope of firms to be investigated, resulting in cost savings on the research side as only a subset is considered for further, more nuanced examination. As already discussed, the model's evaluation was conducted on the Dow Jones Industrial Average through various metrics and statistical tests to be detailed further below. There are five primary categories of results, with specific findings within each sub-category.

CEU eTD Collection

6.2 *Confusion matrix analysis:*

As discussed in the metrics part, the model successfully identifies true positives and demonstrates its robustness in detecting "clear-cut" cases of greenwashing. These are usually companies with large-scale media coverage. Examples include companies like Coca-Cola, Apple, and Amazon. This capability is vital for potential stakeholders who may use the model, such as investors, regulators, and activists who need reliable tools to pinpoint overt greenwashing practices. It also suggests that the model effectively leverages specific features or patterns commonly present in the disclosures of well-documented companies.

On the other hand, the considerable number of false negatives highlights a critical shortfall in the model's sensitivity. Examples of such outcomes include companies like Visa, IBM, and McDonald's. Here, the model does not pick up these companies, as they (according to current theory) use subtler or less overt greenwashing tactics that do not match the more "blatant" patterns the model is set up to detect. This is a possible tactic on the firm side that involves nuanced language in the reporting of already sophisticated CSR initiatives, set up cleverly to superficially address environmental concerns while failing to enact meaningful change. Such avoidance successfully confuses the model, albeit not because these firms directly target NLP-based detection systems. It has more to do with the fact that the model's training corpora lacks the data required to make these nuanced distinctions (there is currently no industry standard pre-training model specialized on greenwashing). This gap can be enhanced by training the model on a broader spectrum of greenwashing expressions and incorporating more diverse data sources that capture these subtler nuances. The latter would require the custom fine-tuning of the BERT model used.

CEU eTD Collection

Moving on, the presence of an identified false positive, Travelers Companies Inc., raises ethical and reputational concerns. Although there is only one such case within the thirty companies, incorrectly labeling a company as engaging in greenwashing can have severe implications. For example, public disclosure of such “results” can damage a firm’s reputation. This underlines the importance of not only refining the model to minimize false positives but also establishing validation mechanisms to confirm greenwashing incidents before public disclosure (more on this in the limitations section). Enhancing the model's precision further can help avoid the consequences that arise from such errors. It also highlights the importance of human validation, as currently, we need a way of creating models that are 100% reliable (and it is questionable if we ever will).

Finally, correctly identifying true negatives, or non-greenwashing companies, such as Salesforce and Intel, reassures users that the model does not wrongly accuse firms of unethical environmental practices. This reliability is crucial if the potential user wishes to maintain the trust of those corporates scrutinized and of the stakeholders as well, who rely on accurate data to make investment and policy decisions. The models' capability to accurately identify the firms here suggests that they are well calibrated to recognize genuine sustainability efforts, which are essential in promoting true corporate responsibility and helping firms committed to environmental stewardship avoid being unfairly penalized.

CEU eTD Collection

6.3 Performance metrics:

Firstly, the most important metric, the model's overall accuracy score, came out to 66.6%, suggesting that two-thirds of the model's predictions align with the actual greenwashing status of companies, which is moderately satisfactory but definitely indicates room for improvement. The critical takeaway is that the model predicts at a rate above chance, indicating that the methodology is valid. The second metric, the precision rate, stands at 91.7%, considered relatively high and essential when deciding if a given model is worth pursuing. Here, it reflects the model's reliability in the cases it identifies as greenwashing, which is necessary for avoiding baseless accusations. This means that whenever the model "decides" a company is guilty of greenwashing, it does so with high accuracy, which is relevant, as anyone implementing such a method must be cautious about levied accusations. The third metric, however, pulls the model's overall standing back a little, as the recall rate is problematic, as over half of actual greenwashing cases, 55%, go undetected. This can be remedied by adjusting the cutoff rate; however, it may also require systematically tuning the embedding BERT model. Lastly, the F1 Score of 68.7% is "decent" and balances precision and recall. However, it confirms the need for enhanced model training to better capture varied greenwashing behaviors.

6.4 Chi-square test:

A Chi-squared test was run on the outputs of the greenwashing detection model to verify its significance, which is needed to see if the model actually “works” or not. The p-value of 0.006 yielded by the test is statistically significant (by an exceedingly high margin). It confirms that the model's predictions are not due to chance and that there is a meaningful association between the predicted and actual cases of greenwashing. This significant result supports the model's effectiveness. It emphasizes the importance of further refinement to address the missed cases rated by the recall metric. Enhancing the model's sensitivity, as mentioned in the previous section, could

CEU eTD Collection

lead to more comprehensive detection capabilities, reducing the risk of type II errors (false negatives).

6.5 Performance across different greenwashing approaches:

The model's performance varies across different greenwashing approaches, highlighting its strengths and limitations. Selective disclosure is detected with what can be considered high accuracy, signaling the model's capability to identify overt misleading claims. Tactics involving misleading communication and decoupling also show good precision and accuracy, suggesting that the pre-trained BERT model well identifies these strategies. On the other hand, however, deception and misinformation are less reliably detected, as indicated by the lower recall rate and higher false negatives, suggesting these areas may require more nuanced or extensive training samples to improve detection. The latter again ties back to the notion of fine-tuning the model which may boost its performance over the metrics mentioned, overall, however, considering the BERT model used is the “base” variant, these results are better than excepted and highlight the strength of transformers-based models.

6.6 Industry analysis

The Dow Jones Index includes an industry component, making it easier to compare the impact on different industries in one go, as opposed to manually having the search for the respective industry of each firm. The analysis of greenwashing potential across industries reveals significant variations: information technology, financial services, and beverage industries are the

CEU eTD Collection

most prone to greenwashing. This could reflect these sectors' intense scrutiny and market pressures regarding environmental claims. On the other hand, industries like healthcare, pharmaceuticals, construction, and mining show lower greenwashing potential, possibly due to different regulatory environments or less direct consumer engagement. Understanding these industry-specific patterns can help tailor the model to be more sensitive to the contextual factors influencing greenwashing behaviors.

6.7 Contributions to the field(s)

The research contributes to two scientific fields: environmental sciences and computer science.Although the main problem it discusses, greenwashing, is mostly within environmental sciences, the proposed framework offers new insights into computer sciences and artificial intelligence.

The initial motivation behind this research was the belief that if AI (NLP) models could better detect and flag greenwashing activities of companies, there would be an increase in transparency, and companies would invest more significant amounts in sustainability initiatives instead of forging numbers and messaging misleading claims. Although the model does not provide a foolproof solution for greenwashing detection, it is undoubtedly an enabler of transparency that could be developed by creating further iterations. Those who wish to replicate this research with a broader range of companies and more extensive data could use the insights to benchmark the sustainability scores of companies against industry standards and identify areas for improvement. The ideal outcome is if corporations are willing to adopt diverse practices instead of their current greenwashing deceptions. Still, the lack of motivation for actual improvement and the more than concerning motivation to put resources into deliberately misguiding customers is

CEU eTD Collection

outside the research's scope. However, as consumers become increasingly aware of differentiating between genuine and manipulative corporate sustainability, there is hope that their trust in corporate sustainability can be restored. More minor attempts like this research serve as a means of restoration in this process.

Although this is not the first research where NLP is used to analyze and interpret textual greenwashing data, it does demonstrate practical applications in real-world greenwashing problems in a new way. The combination of traditional keyword search and NLP used in this thesis is a hybrid approach that leverages both strengths. Although the data of only thirty companies is analyzed, the model does have the potential to explore datasets from more sources (not only Google News) and demonstrates the potential in big data analytics (because with quality data, BERT can handle billions of rows, which counts as "big data") (De Mauro, Greco, and Grimaldi 2016).

AI is not a new topic in environmental sciences, machine learning and deep learning algorithms have been widely discussed and used as attempts to address several environmental challenges. This technology's potential receives praise for helping society manage significant environmental risks and improve environmental safety. Maganathan et al. (2020, 2) best summarize the ecological challenges that AI can potentially address:

> *"...identifying the proper approach for building a framework..., …lack of knowledge in figuring out environmental issues…, …robust way in dealing with environmental data…, …optimal techniques to extract information…, …protocols to prefer the right method for analysis…, information exchange and preserving of data…, …storage cost and energy consumption…".*

CEU eTD Collection

This thesis utilizes NLP techniques, which are different from existing studies that primarily rely on more straightforward methods like keyword matching or content analysis. The challenges are turned into opportunities for greenwashing detection using the NLP components of the suggested model, as they can be flipped to identify possible red flags for greenwashing. Although companies might exploit the lack of knowledge on environmental issues and use complex, ultimately meaningless environmental jargon in their communication and reports, analysis with the research's model can identify such irrelevant or misleading terminology because of context and sentiment processing. Moreover, as already discussed, greenwashing increasingly involves manipulating and selectively presenting data. The proposed framework of the thesis proves to be efficient in assessing data quality and identifying inconsistencies or outliers that indicate the presence of manipulated data.

It is important to note that the contribution of the model, like most NLP models used in environmental sciences, is indirect. It is not directly addressing today's environmental challenges. Still, it helps identify misleading or even false information that can further slow the already stagnating progress toward ecological sustainability and SDGs. Additionally, it is acknowledged that environmental science is broad, and this research focuses only on identifying corporate greenwashing. This focus is further narrowed down as the analyzed articles are from Google News only. Recognizing that greenwashing can occur in various settings, such as government policy or the sponsorship of environmental non-governmental groups is also essential, and the research acknowledges that greenwashing is not limited to business practices.

The presented model's following two impacts are selected as critical contributions to the field of environmental sciences: improved transparency and accountability in greenwashing detection and prioritization of beneficial environment efforts. Transparency and accountability are

CEU eTD Collection

current issues in greenwashing detection. Greenwashing thrives in an environment where the assessment of truthfulness is complex. This lack of transparency and accountability leads to several problems, including misinformed consumers, erosion of trust, and even unfair advantage for greenwashers because they can go unchecked (a detailed explanation can be found in the literature review). Traditional methods for greenwashing detection often fall short because keyword-matching methods rely on identifying specific keywords associated with sustainability. However, companies can quickly adapt their language to bypass such filters. Content analysis methods might analyze a company report's overall sentiment or topics. Still, they lack the sophistication to understand the nuances of language used for greenwashing. Moreover, existing research might have yet to incorporate external data sources for greenwashing detection, such as news articles. Overall, since this research integrates more methodologies, including BERT and cosine similarity, it helps identify deceptive greenwashing tactics that might "escape" more straightforward methods.

Consumers and policymakers misled by greenwashing might prioritize or invest in companies with poor environmental practices, which diverts resources away from genuinely sustainable initiatives with a more significant positive impact. Policymakers ideally would rely on the most accurate possible information about corporate environmental practices to design effective regulations. Greenwashing hinders this process, potentially leading to the creation of policies that fail to address the most pressing environmental issues. Consumers seeking to support environmentally responsible choices are continuously being misled by greenwashing. This wasted attention and spending power slows down progress towards genuine sustainability goals. Traditional computing methods often fall short in these aspects, too. Using keyword matching, content analysis, and limited data sources let greenwashers bypass these methods, as they lack the sophistication to understand the nuances of language used for greenwashing (Lamar 2023).

CEU eTD Collection

However, NLP-based methods can pave the way for standardized and objective greenwashing. According to Kim et al. (2023), this is due to the scalability and automation capabilities of these models, while others like Saranya and Subhashini (2023) think that the reduced human bias will pave the way for a wider use of them.

Analyzing whole articles would lead to an even better understanding of how the media portrays a company's environmental practices. However, the current level of AI greenwashing detection is not at a level to effectively detect multi-layered greenwashing from a couple of sentences, let alone whole articles (Moodaley, Wayne, and Telukdarie 2023). If such sophisticated AI models for full-text analysis of environmental publications already existed, we would see their widespread use. Attempts like EY's "Guide Against Greenwashing Compass" motivate firms to input their text on the website to see sentences that AI detected as greenwashing. Still, even EY admits that their tool has limitations (EY, 2023).

In this research, titles were deliberately chosen as inputs mainly because a model like this one is a precursor to those that will be able to deal with longer texts and must precisely operate on titles (short texts) before processing can be expanded to longer texts. Titles, in general, use concise language, which NLP models can analyze and identify easier than complex sentence structures of a complete article. Additionally, analyzing articles can be computationally and financially resource-intensive, especially when working with large datasets. Concentrating on titles reduces the data volume, speeds up the model's operation, and frees up time for necessary modifications and iterations. While using titles only has limitations, mainly limited context, the focused statements in the titles can improve efficiency and highlight red flags. Titles with phrases like "eco-friendly solution" or "sustainable breakthrough" surely warrant further investigation with the NLP model of the thesis. Focusing solely on titles neglects potentially relevant information from

CEU eTD Collection

analyzing an entire article. This is acknowledged, and a more detailed discussion of limitations is provided in the next section. However, this title-based approach can be a stepping stone towards developing future models that can manage longer texts with even greater accuracy.

6.8 Comparison with keyword matching, the "rule of thumb" detection method

Keyword search detection methods are featured in most academic research investigating greenwashing detection, because these are the current best practice. Assmann and Peralta (2023) developed an NLP-based service that uses keyword search to find corporate greenwashing trends and provides users with indicators for making decisions based on the data they have gathered. In their 2017 study, Griese, Werner, and Hogg delved into event marketing and used keyword searches to identify cases of greenwashing. They observed that despite several detection methods, more thorough methods are needed to address this issue. Pendse, Nerlekar, and Darda (2022) conducted a bibliometric analysis with data from 1996 to 2021, employing keyword search to find trends and patterns in greenwashing practices. These studies agree that keyword search approaches help us better recognize corporate greenwashing, but more advanced methods and tools need to be developed.

Only a limited body of literature explores the application of NLP techniques to identify corporations' greenwashing practices. Makarenko (2023) used content analysis and the Partial Least Squares Structural Equation Modeling (PLS-SEM) method to identify instances of greenwashing in the sustainability reports of Ukrainian agricultural companies. The study by Gutierrez-Bustamante et al. (2022) examined Nordic publicly traded firms and employed natural language processing (NLP) methods to evaluate how well their reports adhere to the Global Reporting Initiative (GRI) framework. Moodaley et al. (2023) also conducted a systematic

CEU eTD Collection

literature review, highlighting the new but promising intersection of AI and greenwashing detection in sustainability reporting. In 2023, Oppong-Tawiah et al. focused on detecting greenwashing on Twitter. They proposed an approach to recognize these misleading tactics based on linguistic signals automatically. The studies by these authors highlight a shift towards more sophisticated methods for greenwashing detection.

While keyword matching offers simplicity when attempting to detect greenwashing, it does have severe limitations. For example, it can result in false positives (flagging non-greenwashing text) and negatives (missing actual greenwashing) since it relies on specific keywords only (Stammbach et al. 2022). This research's framework addresses this issue by utilizing semantic similarity, and based on the results, it does lessen the number of false negatives and positives. The model mainly performs well in terms of false positives, and there was only one company incorrectly flagged as a greenwasher.

Keyword matching may also struggle to react to changing greenwashing strategies as many companies adjust their wording. The model was developed with this in mind, and more approaches were incorporated, making it versatile. The results confirm this versatility. Regarding precision, the rates are high for all approaches (over 88%), which indicates that the model effectively identifies greenwashing across all approaches. The accuracy is also consistent for different methods (67% or 77%), showing the model's ability to maintain its performance. Recall results vary and could be improved (possibly by using a larger dataset with more sources). However, selective disclosure, decoupling, and misleading communication have high recall rates of 75%, 70%, and 70%, respectively, indicating the model's capability to identify these particular tactics.

Implementing keyword matching is generally more straightforward, but it necessitates regular updates to the keyword list. The suggested framework involves meticulous feature

CEU eTD Collection

selection and weighting, ensuring a more effective and adaptable solution. The precision of the model is significantly high (91.7%), indicating that when it identifies greenwashing, it is correct almost every time. Using keyword matching only is not enough because, as already discussed, it often results in higher false positives. This is because it cannot recognize the context in which a keyword is used, leading to incorrect identifications.
The statistical validation provided by the chi-square test confirms that the model's predictions are not based on random chance. This kind of validation would not be possible with simple keyword matching, which has a binary result (a document either matches the keyword or does not) and does not provide enough categories.

To conclude, while research on greenwashing detection usually involves keyword matching, the “feature-enhanced” model of the thesis offers a more sophisticated approach. The computations utilize semantic similarity and predefined definitions from literature to understand the overall sentiment of greenwashing language. This approach has the potential to produce more precise and adaptable results.

6.9 Limitations and practical implications for future research

The thesis results show the proposed model's strengths and limitations. As the research's conclusion draws near, it is crucial to assess its practical implications but acknowledge any limitations that arise. This section aims to provide a roadmap for future enhancements and applications. Here, the specific limitations are discussed, as well as potential real-world impacts that these findings could have on corporate practices, consumer behavior (and decision-making), and potentially even policymaking. The aim is to bridge the gap between academy and practicality and offer insights that could enhance corporate transparency by

CEU eTD Collection

providing more reliable information.

First, the model achieved a precision rate of 91.7%, which is a high accuracy rate for identifying true greenwashing incidents when they are indeed present. However, the recall rate was significantly lower at 55%, which means that almost half of actual greenwashing activities went undetected. Recall is a critical area that must be improved to avoid overlooking greenwashing. Recommendations to do this include rebalancing the dataset and optimizing the F2 score to prioritize recall over precision (Gaurav, 2023).

Next, the variance of effectiveness based on the type of greenwashing approach is a challenge. Specifically, as already mentioned, the model identified selective disclosure with greater accuracy compared to deception and misinformation, which had notably lower recall values. This variation suggests that specific greenwashing strategies are inherently more complex to detect using current NLP methods. This is due to the subtler use of linguistics and less apparent misrepresentations. This limitation shows that the model must be improved to have better linguistic analysis capabilities and to capture a broader range of greenwashing practices.

Additionally, the model performance analysis across different industries revealed varying levels of successful detection. Sectors such as information technology and financial services showed higher instances of detected greenwashing due to their more frequent and publicized sustainability claims. However, industries like healthcare and pharmaceuticals showed lower detection rates, which could impact the applicability and reliability of the model across sectors. This industry-specific variance urges the creation of more customized NLP models that can catch the unique linguistic and operational characteristics of each industry. Only this customization could ensure a more uniform application in the long run.

CEU eTD Collection

The high precision of the model could offer a monitoring method for environmental advocacy. Delving into the different greenwashing approaches and creating specific keywords for each could help monitor and verify corporate sustainability claims more effectively. This enhanced monitoring would be handy since not all greenwashing cases are the same, and specific practices might require different retorsions than others. Even though the model was used on 30 companies only in this research, mainly due to time constraints, the results already show that specific industries are at high risk for greenwashing detection. An improved model could be beneficial for creating targeted policies that specifically address those greenwashing techniques that are particularly prevalent in these high-risk industries. Eventually, policymakers could design standards that demand higher corporate transparency by understanding which sectors are more inclined to greenwash and which are not as much.

When considering future research based on this thesis, there are avenues to expand and refine the study, potentially enhancing the model's robustness, applicability, and impact. As the literature review demonstrates, the fields of machine learning and NLP are not new researchers have actively been studying different models for around 80 years. However, while weaker computational power slowed progress decades ago, the computational power to train top AI models has doubled every three and a half months since 2012 (Hao, 2019). At the same time, so-called quantization (the reduction of the size of large language models without sacrificing much of their precision) has enabled researchers to train and run top-tier models locally from a personal computer (Li et al. 2024). As AI models and computational capabilities continue to evolve, the integration of full-text analysis alongside title-based detection can be explored. This could involve techniques like sentiment analysis of the entire article. As previously said, processing full-length

CEU eTD Collection

articles would require more resources, both financially and in terms of computer capability. However, once these hurdles are overcome, the potential benefits will be significant.

Starting with the first recommendation, the dataset could be expanded to include data from sources that are not written in English to capture a broader spectrum of articles, potentially, different linguistic and cultural contexts. Currently, the research is localized to the US since Dow Jones companies were used to evaluate the framework. Adding more countries to the research would improve the global applicability of the model and could be used in diverse markets. Besides adding articles from more countries, more industries should be added, too. An analysis of the differences between results in different sectors is demonstrated in the results section, but there are still industries that have not been introduced to the model. Broadening the dataset to include more industries, especially those underrepresented in this research, could test the model's effectiveness across more sectors and identify industry-specific greenwashing patterns.

Additionally, as discussed, the field of NLP is a fast-advancing one, and although BERT is a recent model, newer transformer-based models exist beyond BERT (such as GPT-3). Incorporating one of these or combining more models could offer improved, up-to-date context understanding and semantic analysis capabilities. Speaking of updating, there exist methods for continuous learning, where the model periodically updates itself with new data. For example, ensemble methods could be added to BERT to create more models working on the same problem. Incorporating these could ensure adaptation to the evolving language and greenwashing tactics, always offering up-to-date information and results (Jia, Wen, and Youzhi 2023).

In this research, the model *is not a regression model*. The test data is the "real" data (whether there has been greenwashing or not), and the predictions were the model's predictions. We have actual data and predicted data, not prediction and evaluation. Since the model is

CEU eTD Collection

prediction agnostic, it can be fitted to any other dataset (for example, Nasdaq, S&P 500, with parameter tweaking). The only place when an individual using this framework would need to use testing would be if they were either comparing the performance of two different embedding models or if they were comparing the performance of other fine tunes of the same embedding model. In the research's case, one embedding model was taken, the industry-standard BERT; the rest of the prediction model is fixed and non-varying. Unlike regression models, the predictions of this model do not involve weights in the traditional sense. The only variability can come from the change mentioned above of the embedding model component. However, it is acknowledged that more rigorous cross-validation techniques could be used in future research to ensure the model's reliability and generalizability across different datasets and external validation sets.

Another suggestion would be to improve the model's explainability. Most NLP models are black boxes, meaning that the user is most likely to see only the output, not how exactly the model reached its conclusions (Hassija et al. 2024). BERT uses complex algorithms to analyze text and assign "greenwashing scores" to companies. Its algorithms involve layers of interconnected mathematical functions that are difficult to interpret for humans. The user is faced with the fact that they are looking at an output, yet the model's inner workings cannot be seen.

The more explainable models (white-box models like decision trees and linear regressions) are more accessible to interpret and understand, as they allow us to see how the model arrives at its results. However, they are typically also less accurate than black-box models and may struggle with complex data (Erico and Cuntai 2019). While researchers are developing techniques to make black-box models more interpretable, in the current AI research space as of mid-2024, black-box models, particularly deep learning models, are more prominent than white-box models (the best example being the most popular deep learning model, ChatGPT, which is in fact a black-box).

CEU eTD Collection

Their dominance is due mainly to advancements in hardware, computational power, their focus on accuracy, and data availability. However, in spaces where trust and understanding are crucial, such as healthcare, there is a push for explainability (Metta et al. 2024).

The last suggestion for future research is to consider policy and regulatory implications. The research touched upon greenwashing as a violation of regulations concerning fair advertising and consumer protection. However, policy recommendations are not offered due to time constraints. A model designed to identify greenwashing practices needs to be aware of related regulations to ensure its outputs align with the legal description of greenwashing. For instance, a model might flag a company's statement as greenwashing based on technicalities that would not hold up in court. An enhanced model would stay up-to-date on policy developments that maintain accuracy and relevance. Demonstrating a commitment to responsible AI development would also be crucial. Some directions regarding policy where future research could start include identifying regulatory gaps to see practices that are not currently captured by existing greenwashing regulations, the standardization of greenwashing definitions, and the establishment of more effective enforcement rules.

These were the points that the research could not tackle or did not consider. These limits do offer some exciting avenues for future research. By incorporating multiple data sources, advanced text analysis techniques, and a self-reflective learning approach, the model can evolve into a powerful sentence similarity model used against corporate greenwashing.

CEU eTD Collection

7. Chapter: Conclusion

This research seeks to validate current NLP techniques' applicability to the intangible greenwashing area. Predicting greenwashing is difficult, not only due to the deliberate deception by companies that invest significant resources into concealing it, but also because there are no established regulations to clearly define what greenwashing is. The topic of greenwashing is usually tackled from the financial side, for example, auditors comb through the financial records of firms for evidence to prove that they are, in fact, less green than they claim to be. This is effective, but it is also costly and immensely time-consuming. Without a valid will on the side of governments to tackle the issue, firms will continue to get around most current and proposed regulations.

The thesis addresses greenwashing through an analysis of semantics - specifically, examining companies' messaging and translating it into numerical data to identify potential deceptive practices. A couple of years ago, this may not have been possible. However, with the development of the Transformers machine learning architecture, we can now capture semantic meaning in numbers, which consequently allows for the numerical comparison of specific phrases with other phrases. This is where the core idea of the research lies. The text compares a set of five comprehensive definitions of greenwashing to a collection of news headlines taken from Google News. The comparison is designed to give higher scores to titles more likely suggesting the mentioned practices.

As expected, the firms aim to curb their claims by deliberately using language that makes their environmental impact minor. This is why a second metric, a ranking, was introduced. This frequency is the count of articles related to a particular firm. The architecture consists of the

CEU eTD Collection

average semantic greenwashing scores for all articles about a given firm multiplied by the number of articles. Each company can be assigned a "greenwashing score" based on my articles, using a specific formula of more methods working together. This approach means that a company with a lower chance of greenwashing, but more articles written about it could be ranked lower than a company with fewer articles but a higher probability of greenwashing.

This framework was validated on the Dow Jones Industrial Average. It is shown to predict at a higher level than random by a large margin (0.006 as compared to the p-value method's 0.05). Therefore, it can be determined that this approach is empirically valid. This can mean a much cheaper way of scanning, evaluating, and finally predicting greenwashing in a few firms and over entire industries at a much more affordable rate than most other methods would allow for. There are also multiple ways to extend the research's framework, as it can be appended by different news sources, lists, and even assisting models used for the backend architecture.

As the methodology is shown to be one that can predict greenwashing, attempts at improvement are not only welcome but actively encouraged. Through the power of collaborative, open-source research, we can make an impact against the deceptions unfortunately still woven through the fabric of efforts towards a greener economy.

CEU eTD Collection

8. Appendix

8.1 *Appendix A: Code Repository on Github*

The code used throughout the research is publicly available on GitHub. It is shared under the Apache License 2.0, allowing for free use, modification, and distribution. The repository includes detailed instructions for replicating the results, and can be accessed at the following URL:

GitHub Repository Link

CEU eTD Collection

8.2 *Appendix B: Code Output*

This appendix shows data, which is the output of the code, from which, part of the results are deducted. It is extracted from an Excel file and has been divided into three sections for fit within the document. Table 5 lists each company alongside their respective scores for the five greenwashing approaches analyzed. Table 6 ranks the companies according to each approach, with rankings adjusted based on the number of articles to weigh relevance (this is discussed in detail in the Model Architecture section). Table 7 presents the model's predictions for each company, where "0" indicates no greenwashing and "1" indicates presence of greenwashing, including data on confirmed cases.

Company Name	Average General_Definition Score	Average Deception_and_Misinformation Scor	Average Misleading_Communication Scor	Average Selective_Disclosure Score	Average Greenwashing_as_Decoupling Scor
3M greenwashing	0,231	0,234	0,264	0,231	0,237
Amazon greenwashing	0,416	0,313	0,332	0,285	0,342
American Express greenwash	0,364	0,280	0,335	0,288	0,346
Amgen greenwashing	0,241	0,255	0,284	0,273	0,245
Apple greenwashing	0,412	0,327	0,338	0,295	0,354
Boeing greenwashing	0,347	0,253	0,285	0,245	0,305
Caterpillar greenwashing	0,202	0,187	0,242	0,216	0,252
Chevron greenwashing	0,409	0,326	0,342	0,331	0,352
Cisco greenwashing	0,332	0,257	0,312	0,291	0,325
Coca-Cola greenwashing	0,509	0,377	0,380	0,312	0,399
Disney greenwashing	0,341	0,290	0,308	0,279	0,331
Dow greenwashing	0,316	0,279	0,314	0,276	0,300
Goldman Sachs greenwashing	0,241	0,209	0,210	0,213	0,219
Home Depot greenwashing	0,303	0,247	0,319	0,270	0,306
Honeywell greenwashing	0,237	0,201	0,286	0,264	0,270
IBM greenwashing	0,312	0,253	0,317	0,285	0,315
Intel greenwashing	0,268	0,224	0,288	0,255	0,269
Johnson & Johnson greenwas	0,364	0,274	0,309	0,269	0,340
JP Morgan Chase greenwashi	0,344	0,251	0,285	0,272	0,287
McDonalds greenwashing	0,375	0,309	0,345	0,287	0,359
Merck greenwashing	0,194	0,263	0,325	0,271	0,268
Microsoft greenwashing	0,454	0,331	0,387	0,338	0,391
Nike greenwashing	0,425	0,352	0,326	0,260	0,373
Procter & Gamble greenwash	0,488	0,335	0,346	0,301	0,372
Salesforce greenwashing	0,311	0,236	0,319	0,307	0,315
Travelers greenwashing	0,411	0,270	0,328	0,291	0,376
UnitedHealth Group greenwa	0,321	0,317	0,396	0,372	0,389
Verizon greenwashing	0,267	0,215	0,252	0,256	0,258
Visa greenwashing	0,365	0,214	0,244	0,215	0,270
Walmart greenwashing	0,403	0,334	0,353	0,289	0,352

Table 5: Table showing the code's output, each company's "greenwashing score" by greenwashing approach

Company Name	Number of Articles	General_Definition Ranking	Deception_and_Misinformation Ranking	Misleading_Communication Ranking	Selective_Disclosure Ranking	Greenwashing_as_Decoupling Ranking
3M greenwashing	29	6,685	6,800	7,644	6,709	6,860
Amazon greenwashing	60	24,964	18,761	19,897	17,110	20,520
American Express greenwashing	32	11,641	8,947	10,705	9,222	11,083
Amgen greenwashing	2	0,482	0,509	0,569	0,546	0,491
Apple greenwashing	63	25,951	20,607	21,318	18,604	22,307
Boeing greenwashing	51	17,701	12,879	14,525	12,484	15,551
Caterpillar greenwashing	21	4,242	3,937	5,089	4,530	5,285
Chevron greenwashing	49	20,064	15,966	16,740	16,236	17,233
Cisco greenwashing	31	10,280	7,981	9,684	9,032	10,083
Coca-Cola greenwashing	98	49,889	36,984	37,203	30,574	39,120
Disney greenwashing	26	8,877	7,553	8,019	7,241	8,601
Dow greenwashing	51	16,140	14,240	16,013	14,095	15,316
Goldman Sachs greenwashing	39	9,396	8,140	8,196	8,308	8,558
Home Depot greenwashing	45	13,629	11,107	14,343	12,133	13,754
Honeywell greenwashing	17	4,023	3,411	4,855	4,483	4,586
IBM greenwashing	38	11,850	9,598	12,036	10,837	11,980
Intel greenwashing	43	11,536	9,632	12,382	10,965	11,553
Johnson & Johnson greenwashing	42	15,294	11,511	12,991	11,302	14,268
JP Morgan Chase greenwashing	46	15,823	11,530	13,125	12,530	13,208
McDonalds greenwashing	24	8,994	7,409	8,286	6,894	8,618
Merck greenwashing	2	0,388	0,525	0,651	0,543	0,535
Microsoft greenwashing	46	20,868	15,231	17,823	15,561	18,005
Nike greenwashing	37	15,730	13,020	12,074	9,603	13,795
Procter & Gamble greenwashing	29	14,144	9,728	10,029	8,720	10,799
Salesforce greenwashing	42	13,046	9,919	13,385	12,881	13,217
Travelers greenwashing	52	21,380	14,039	17,073	15,143	19,536
UnitedHealth Group greenwashing	4	1,285	1,267	1,585	1,487	1,555
Verizon greenwashing	29	7,735	6,245	7,294	7,414	7,492
Visa greenwashing	40	14,584	8,548	9,757	8,586	10,816
Walmart greenwashing	43	17,328	14,360	15,179	12,413	15,132

Table 6: Ranking of each company in each approach

CEU eTD Collection

Company Name	Model prediction	Confirmed Greenwashing	Comfirmed Greenwashing	Industry
3M greenwashing	0	no		Conglomerate
Amazon greenwashing	1	yes	https://news.google.com/articles/	Retailing
American Express greenwashing	0	no		Financial Services
Amgen greenwashing	0	no		Biopharmaceutical
Apple greenwashing	1	yes	https://news.google.com/articles/	Information Technology
Boeing greenwashing	1	yes	https://www.citywatchla.com/los-	Aerospace and Defense
Caterpillar greenwashing	0	yes	https://www.theguardian.com/su	Construction and Mining
Chevron greenwashing	1	yes	https://news.google.com/articles/	Petroleum Industry
Cisco greenwashing	0	no		Information Technology
Coca-Cola greenwashing	1	yes	https://news.google.com/articles/	Beverage Industry
Disney greenwashing	0	yes	https://news.google.com/articles/	Broadcasting and Entertainment
Dow greenwashing	1	yes	https://news.google.com/articles/	Chemical Industry
Goldman Sachs greenwashing	0	yes	https://news.google.com/articles/	Financial Services
Home Depot greenwashing	0	yes	https://news.google.com/articles/	Home Improvement
Honeywell greenwashing	0	no		Conglomerate
IBM greenwashing	0	yes	https://www.dailymail.co.uk/scier	Information Technology
Intel greenwashing	0	no		Semiconductor Industry
Johnson & Johnson greenwashing	1	yes	https://palmoildetectives.com/20	Pharmaceutical Industry
JP Morgan Chase greenwashing	1	yes	https://news.google.com/articles/	Financial Services
McDonalds greenwashing	0	yes	https://news.google.com/articles/	Food Industry
Merck greenwashing	0	no		Pharmaceutical Industry
Microsoft greenwashing	1	yes	https://www.ethicalconsumer.org	Information Technology
Nike greenwashing	1	yes	https://news.google.com/articles/	Apparel
Procter & Gamble greenwashing	0	yes	https://news.google.com/articles/	Consumer Goods
Salesforce greenwashing	0	no		Information Technology
Travelers greenwashing	1	no		Insurance
UnitedHealth Group greenwashing	0	no		Health Care
Verizon greenwashing	0	yes	https://news.google.com/articles/	Telecommunications
Visa greenwashing	0	yes	https://news.google.com/articles/	Financial Services
Walmart greenwashing	1	yes	https://news.google.com/articles/	Retailing

Table 7: Companies' model prediction vs confirmed greenwashing (model prediction is true/1 if greenwashing score is above the threshold)

CEU eTD Collection

9. Bibliography

6Clicks. "What Are the 3 Pillars of ESG?" Accessed February 20, 2024. https://www.6clicks.com/resources/answers/what-are-the-3-pillars-of-esg#:~:text=ESG%20stands%20for%20environmental%2C%20social,the%20depletion%20of%20natural%20resources.

Akyildirim, E., S. Corbet, S. Ongena, and L. Oxley. 2023. "Greenwashing: Do Investors, Markets and Boards Really Care?" *Swiss Finance Institute Research Paper*, no. 23-90.

Amenc, Noel, Lionel Martellini, and Frank J. Fabozzi. 2021. "Doing Good or Feeling Good?" August.

Architect Magazine. "Study: 95% of Green-Marketed Products Utilize Misleading Claims." Accessed March 22, 2024. https://www.architectmagazine.com/technology/products/study-95-of-green-marketed-products-utilize-misleading-claims_o.

Arize. "F1 Score: What It Is and Why It Matters." Accessed May 9, 2024. https://arize.com/blog-course/f1-score/#:~:text=F1%20score%20is%20a%20measure,can%20be%20modified%20into%20F0.

arXiv. "On the Effectiveness of Contextual Embeddings for Answer Sentence Selection." *arXiv preprint*. https://arxiv.org/abs/1907.07374.

Atkins, Margaret, and Robin Osborne, eds. 2006. *Poverty in the Roman World*. Cambridge: Cambridge University Press.

BARRÉ, Rémi, ed. 2003. *Diasporas Scientifiques: Comment les Pays en Développement Peuvent-Ils Tirer Parti de Leurs Chercheurs et de Leurs Ingénieurs Expatriés?* New edition [online]. Marseille: IRD Éditions. Accessed March 5, 2024. http://books.openedition.org/irdeditions/2591. ISBN: 978-2-7099-1770-4. https://doi.org/10.4000/books.irdeditions.2591.

Bawden, David, and Lyn Robinson. "Information Overload: An Introduction." *Oxford Research Encyclopedia of Politics*. June 30, 2020. Accessed March 7, 2024. https://oxfordre.com/politics/view/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-1360.

Berkson, J. 1942. "Tests of Significance Considered as Evidence." *Journal of the American Statistical Association* 37: 325–335. https://doi.org/10.2307/2279000.

Bernini, F., and F. La Rosa. 2023. "Research in the Greenwashing Field: Concepts, Theories, and Potential Impacts on Economic and Social Value." *Journal of Management & Governance*. https://doi.org/10.1007/s10997-023-09686-5.

CEU eTD Collection

Bingler, Julia, Mathias Kraus, and Markus Leippold. 2021. "Cheap Talk and Cherry-Picking: What ClimateBert Has to Say on Corporate Climate Risk Disclosures." March 2. Available at SSRN: https://ssrn.com/abstract=3796152 or http://dx.doi.org/10.2139/ssrn.3796152.

Bothello, Joel, Ioannis Ioannou, Vlad-Andrei Porumb, and Yasemin Zengin-Karaibrahimoglu. 2023. "CSR Decoupling Within Business Groups and the Risk of Perceived Greenwashing." *Strategic Management Journal*. First published: 02 July 2023. https://doi.org/10.1002/smj.3532.

Bowen, H. R. 1953. *Social Responsibilities of the Businessman*. University of Iowa Press.

Bromley, Jane, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah. 1994. "Signature Verification Using a 'Siamese' Time Delay Neural Network." *Advances in Neural Information Processing Systems* 6: 737–744.

Brouwer, Anne. 2016. "Revealing Greenwashing: A Consumers' Perspective." La Trobe University, Plenty Road & Kingsbury Drive, Melbourne VIC 3086, Australia.

Burke, L., and J. M. Logsdon. 1996. "How Corporate Social Responsibility Pays Off."

Business Leadership Today. "What Are the Three Pillars of a Sustainable Business?" Accessed March 15, 2024. https://businessleadershiptoday.com/what-are-the-three-pillars-of-a-sustainable-business/#:~:text=Social%20Pillar,are%20part%20of%20their%20practices.

Cambridge Dictionary. "Greenwashing." Accessed February 28, 2024. https://dictionary.cambridge.org/dictionary/english/greenwashing.

Carbon Market Watch. *Corporate Climate Responsibility Monitor 2023*. Accessed April 7, 2024. https://carbonmarketwatch.org/wp-content/uploads/2023/02/CorporateClimateResponsibilityMonitor2023.pdf.

Carlson, L., S. J. Grove, and N. Kangun. 1993. "A Content Analysis of Environmental Advertising Claims: A Matrix Method Approach." *Journal of Advertising* 22 (3): 27–39.

Carroll, A. B. 1979. "A Three-Dimensional Conceptual Model of Corporate Performance." *Academy of Management Review* 4 (4): 497–505.

Cleanhub. "Greenwashing Examples." Accessed February 28, 2024. https://blog.cleanhub.com/greenwashing-examples#:~:text=One%20of%20the%20most%20famous,31%20billion%20euros%20%E2%80%94%20so%20far.

CEU eTD Collection

Cochran, P. L., and R. A. Wood. 1984. "Corporate Social Responsibility and Financial Performance." *The Academy of Management Journal* 27 (1): 42–56.

Cojoianu, Theodor, Andreas Hoepner, Georgiana Ifrim, and Yanan Lin. 2020. "Greenwatch-Shing: Using AI to Detect Greenwashing." June.

DataStax. "What Is Cosine Similarity." Accessed May 9, 2024. https://www.datastax.com/guides/what-is-cosine-similarity.

Davis, K. 1960. "Can Business Afford to Ignore Social Responsibilities?" *California Management Review* 2 (3): 70–76.

Davoudi, Leonardo, Christopher McKenna, and Rowena Olegario. 2018. "The Historical Role of the Corporation in Society." *Journal of the British Academy* 6: 17-47. https://doi.org/10.5871/jba/006s1.017.

de Freitas Netto, S.V., M.F.F. Sobral, A.R.B. Ribeiro, et al. 2020. "Concepts and Forms of Greenwashing: A Systematic Review." *Environmental Sciences Europe* 32: 19. https://doi.org/10.1186/s12302-020-0300-3.

de Freitas Netto, S.V., M.F.F. Sobral, A.R.B. Ribeiro, et al. 2020. "Concepts and Forms of Greenwashing: A Systematic Review." *Environmental Sciences Europe* 32: 19. https://doi.org/10.1186/s12302-020-0300-3.

De Mauro, Andrea, Marco Greco, and Michele Grimaldi. 2016. "A Formal Definition of Big Data Based on Its Essential Features." *Library Review* 65: 122–135. https://doi.org/10.1108/LR-06-2015-0061.

DiPietro, Robert, and Gregory D. Hager. 2020. "Deep Learning: RNNs and LSTM." In *Handbook of Medical Image Computing and Computer Assisted Intervention*, edited by S. Kevin Zhou, Daniel Rueckert, and Gabor Fichtinger, 503-519. The Elsevier and MICCAI Society Book Series. Academic Press. ISBN 9780128161760. https://doi.org/10.1016/B978-0-12-816176-0.00026-0.

Doyle, Charles. 2016. *A Dictionary of Marketing*. 4th ed. Oxford: Oxford University Press.

Earth.Org. "Greenwashing: Companies and Corporations." Accessed March 10, 2024. https://earth.org/greenwashing-companies-corporations/.

Elkington, J. 1998. "Partnerships from Cannibals with Forks: The Triple Bottom Line of 21st-Century Business." *Environmental Quality Management* 8 (1): 37–51.

ESMA. "Progress Report: ESMA Response to COM RfI on Greenwashing Risks (EN)." ESMA30-1668416927-2498, Sustainable Finance, Report.

CEU eTD Collection

European Commission. " Press Release." Press Corner. Accessed January 15, 2024. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_1692.

European Commission. "Europe's Moment: Repair and Prepare for the Next Generation." Accessed May 9, 2024. https://circabc.europa.eu/ui/group/44278090-3fae-4515-bcc2-44fd57c1d0d1/library/b11ba10b-5049-4564-b47a-51a9bc9003c8/details?download=true.

European Commission. 2023. "Proposal for a Directive of the European Parliament and of the Council on Substantiation and Communication of Explicit Environmental Claims (Green Claims Directive)." COM(2023) 166 final, 2023/0085(COD), Brussels, 22 March 2023.

European Parliament. "'Green Claims' Directive: Protecting Consumers from Greenwashing." European Parliamentary Research Service. Accessed November 30, 2023. https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2023)753958.

EY. "Can Artificial Intelligence Uncover Greenwashing?" Accessed May 9, 2024. https://www.ey.com/en_no/digital/can-artificial-intelligence-uncover-greenwashing.

Financial Advisor Magazine. "Americans Need to Learn to Live More Like Europeans." Accessed December 22, 2023. https://www.fa-mag.com/news/americans-need-to-learn-to-live-more-like-europeans-64875.html.

Fischer, M., et al. 2023. "Corporate Sustainability." In *Sustainable Business*, SpringerBriefs in Business. Cham: Springer. https://doi.org/10.1007/978-3-031-25397-3_4.

Gräuler, Matthias, and Frank Teuteberg. 2014. "Greenwashing in Online Marketing – Investigating Trust-Building Factors Influencing Greenwashing Detection."

Griese, Kai-Michael, Werner Kim, and Johannes Hogg. 2017. "Avoiding Greenwashing in Event Marketing: An Exploration of Concepts, Literature and Methods." *The Journal of Men's Studies*. doi:10.5539/JMS.V7N4P1.

Gutierrez-Bustamante, Marcelo, and Leonardo Espinosa-Leal. 2022. "Natural Language Processing Methods for Scoring Sustainability Reports—A Study of Nordic Listed Companies." *Sustainability*. doi:10.3390/su14159165.

Hao, Karen. 2019. "The Computing Power Needed to Train AI Is Now Rising Seven Times Faster than Ever Before." *MIT Technology Review*, November 11, 2019. https://www.technologyreview.com/2019/11/11/132004/the-computing-power-needed-to-train-ai-is-now-rising-seven-times-faster-than-ever-before/.

Harris, Zellig S. 1954. "Distributional Structure." *WORD* 10 (2-3): 146-162. https://doi.org/10.1080/00437956.1954.11659520.

CEU eTD Collection

Harris, Zellig S. 1954. "Distributional Structure." *WORD* 10 (2-3): 146-162. https://doi.org/10.1080/00437956.1954.11659520.

Hassija, V., V. Chamola, A. Mahapatra, et al. 2024. "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence." *Cognitive Computation* 16: 45–74. https://doi.org/10.1007/s12559-023-10179-8.

Hugging Face. 2023. "BERT." Accessed February 21, 2024. https://huggingface.co/docs/transformers/en/model_doc/bert

Hugging Face. "Sentence Similarity." Accessed February 21, 2024. https://huggingface.co/tasks/sentence-similarity.

Investopedia. "Chi-Square Statistic." Accessed May 9, 2024. https://www.investopedia.com/terms/c/chi-square-statistic.asp#:~:text=Chi%2Dsquare%20tests%20are%20often,of%20variables%20in%20the%20relationship.

Investopedia. "Three Pillars of Corporate Sustainability." Accessed January 29, 2024. https://www.investopedia.com/articles/investing/100515/three-pillars-corporate-sustainability.asp#:~:text=Corporate%20sustainability%20is%20to%20an,in%20companies%20big%20and%20small.

Jan, Assmann, and Alberto Peralta. 2023. "Establishment of NLP-Based Greenwashing Pattern Detection Service." *Lecture Notes in Electrical Engineering*. https://doi.org/10.1007/978-981-99-1252-0_32.

Jenkins, Rhonda. "Regulating Corporate Greenwashing in the United States: Combating a Growing Champion of Misleading Marketing." Accessed January 10, 2024. https://www.linkedin.com/pulse/regulating-corporate-greenwashing-united-states-combating-champion-o/.

Jia, Jianguo, Wen Liang, and Youzhi Liang. 2023. "A Review of Hybrid and Ensemble in Deep Learning for Natural Language Processing." arXiv. Accessed June 1, 2024. https://arxiv.org/html/2312.05589v1.

Johnson, Lamar. 2023. "Greenwashing Rising: Report Sees Reputational Risk From Social Washing and Sustainabilitywashing." *ESG Dive*, October 11. Accessed June 4, 2024. https://www.esgdive.com/news/greenwashing-rising-report-rep-risk-social-washing-sustainability/696289/.

Jones, T. M. 1980. "Corporate Social Responsibility Revisited, Redefined." *California Management Review* 22 (3): 59–67.

Jung, Jae, and Su Park. 2017. "Volkswagen's Diesel Emissions Scandal." *Thunderbird International Business Review* 59. https://doi.org/10.1002/tie.21876.

CEU eTD Collection

Jurafsky, Daniel, and H. James Martin. 2000. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Upper Saddle River, N.J.: Prentice Hall. ISBN 978-0-13-095069-7.

Kadyan, Jagbir. 2016. "United Nations Global Compact and Corporate Social Responsibility." *International Journal of Science and Research (IJSR)* 5: 478–480.

Kang, Hyewon, and Jinho Kim. 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.

Karadimitriou, A., T. von Krogh, C. Ruggiero, C. Biancalana, M. Bomba, and W. H. Lo. 2022. "Investigative Journalism and the Watchdog Role of News Media: Between Acute Challenges and Exceptional Counterbalances." In *Success and Failure in News Media Performance: Comparative Analysis in the Media for Democracy Monitor 2021*, 101–125. https://doi.org/10.48335/9789188855589-5.

Kassinis, George, and Alexia Panayiotou. 2017. "Visuality as Greenwashing: The Case of BP and Deepwater Horizon." *Organization & Environment*. https://doi.org/10.1177/1086026616687014.

Kim, J.S., J.B. Sim, Y.J. Kim, M.K. Park, S.J. Oh, and I.C. Doo. 2023. "Establishment of NLP-Based Greenwashing Pattern Detection Service." In *Advances in Computer Science and Ubiquitous Computing*, edited by J.S. Park, L.T. Yang, Y. Pan, and J.H. Park, vol. 1028. Singapore: Springer. https://doi.org/10.1007/978-981-99-1252-0_32.

Levenshtein, Vladimir I. 1966. "Binary Codes Capable of Correcting Deletions, Insertions, and Reversals." *Soviet Physics Doklady* 10 (8): 707-710.

Liu, Yeyi, Martin Heinberg, Xuan Huang, and Andreas B. Eisingerich. 2023. "Building a Competitive Advantage Based on Transparency: When and Why Does Transparency Matter for Corporate Social Responsibility?" *Business Horizons* 66 (4): 517-527. https://doi.org/10.1016/j.bushor.2022.10.004.

Locate2u. "Blue Yonder: Consumers Willing to Prioritize Sustainability Over Price and Convenience." Accessed March 1, 2024. https://www.locate2u.com/retail/blue-yonder-consumers-willing-to-prioritize-sustainability-over-price-convenience/.

Lokuwaduge, Chitra S. de Silva, and Keshara Manindri De Silva. 2022. "ESG Risk Disclosure and the Risk of Green Washing." *Australasian Accounting, Business and Finance Journal* 16 (1): 146-159. https://doi.org/10.14453/aabfj.v16i1.11.

Lueg, Klarissa, and Rainer Lueg. 2020. "Detecting Green-Washing or Substantial Organizational Communication: A Model for Testing Two-Way Interaction Between Risk and Sustainability Reporting." *Sustainability* 12. https://doi.org/10.3390/su12062520.

CEU eTD Collection

Maganathan, Tharsanee, et al. 2020. "Analyzing the Leading Causes of Traffic Fatalities Using XGBoost and Grid-Based Analysis: A City Management Perspective." *IOP Conference Series: Materials Science and Engineering* 955: 012107.

Makarenko, Inna. 2023. "Greenwashing Detection and Impact on Responsible Business and Investment: Case of Ukrainian Companies (Agriculture Leaders)." *Springer Proceedings in Business and Economics*, 129-145. doi:10.1007/978-3-031-28131-0_10.

McMillan, Kathryn, and Robert McLure. 2003. "Free/Open Source Software (FOSS): A Research Agenda." *First Monday* 8 (3). https://opensource.org/.

Medium. "Mastering Classification Metrics: A Beginner's Guide Part 2: F1, F0.5, and F2 Scores." Accessed May 9, 2024. https://medium.com/@prateekgaurav/mastering-classification-metrics-a-beginners-guide-part-2-f1-f0-5-and-f2-scores-154d5c729085.

Medium. "The Decline of Print Media: Why It Matters." Accessed May 9, 2024. https://medium.com/illumination/the-decline-of-print-media-why-it-matters-c06181090d14.

Merriam-Webster. "Greenwash." Accessed December 19, 2023. https://www.merriam-webster.com/dictionary/greenwash.

Metta, C., A. Beretta, R. Pellungrini, S. Rinzivillo, and F. Giannotti. 2024. "Towards Transparent Healthcare: Advancing Local Explanation Methods in Explainable Artificial Intelligence." *Bioengineering (Basel)* 11 (4): 369. https://doi.org/10.3390/bioengineering11040369.

Montgomery, A. W., T. P. Lyon, and J. Barg. 2023. "No End in Sight? A Greenwash Review and Research Agenda." *Organization & Environment* 0 (0). https://doi.org/10.1177/10860266231168905.

Moodaley, Wayne, and Arnesh Telukdarie. 2023. "Greenwashing, Sustainability Reporting, and Artificial Intelligence: A Systematic Literature Review." *Sustainability* 15 (2): 1481. https://doi.org/10.3390/su15021481.

Moodaley, Wayne, and Arnesh Telukdarie. 2023. "Greenwashing, Sustainability Reporting, and Artificial Intelligence: A Systematic Literature Review." *Sustainability*. doi:10.3390/su15021481.

Murphy, A.H. 1996. "The Finley Affair: A Signal Event in the History of Forecast Verification." *Weather and Forecasting* 11: 3-20.

Murphy, Allan H. 1996. "The Finley Affair: A Signal Event in the History of Forecast Verification." *Weather and Forecasting* 11 (1): 3. Bibcode:1996WtFor..11....3M. https://doi.org/10.1175/1520-0434(1996)011

CEU eTD Collection

Nguyen, Ngoc. 2023. *Fast Fashion & Greenwashing: The Worst Combination for Sustainability*.

Nilsson, M., D. Griggs, and M. Visbeck. 2016. "Map the Interactions Between Sustainable Development Goals." *Nature* 534: 320–322. https://doi.org/10.1038/534320a.

OECD. 2022. *Building Trust and Reinforcing Democracy: Preparing the Ground for Government Action*. OECD Public Governance Reviews. Paris: OECD Publishing. https://doi.org/10.1787/76972a4a-en.

Offset Guide. "What Is a Carbon Offset?" Accessed January 29, 2024. https://www.offsetguide.org/understanding-carbon-offsets/what-is-a-carbon-offset/.

Open Source Initiative (OSI). "The Definition of Open Source." Accessed May 9, 2024. https://opensource.org/.

OpenAI. "New and Improved Embedding Model." Accessed February 21, 2024. https://openai.com/blog/new-and-improved-embedding-model.

Oppong-Tawiah, Divinus, and Jane Webster. 2023. "Corporate Sustainability Communication as 'Fake News': Firms' Greenwashing on Twitter." *Sustainability*. doi:10.3390/su15086683.

Orange, E., and A. M. Cohen. 2010. "From Eco-Friendly to Eco-Intelligent." *The Futurist* 44 (5): 28–32.

Oxford English Dictionary. "Greenwashing." Accessed April 12, 2024. https://www.oed.com/dictionary/greenwashing_n?tl=true.

Papadimitriou, Christos, Prabhakar Raghavan, Hisao Tamaki, and Santosh Vempala. 1998. "Latent Semantic Indexing." In *Proceedings of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS '98)*, 159-168. https://doi.org/10.1145/275487.275505. ISBN 978-0897919968. Archived from the original (Postscript) on May 9, 2013. Retrieved April 17, 2012.

Pedersen, Esben Rahbek Gjerdrum, Gorm Gabrielsen, Signe Bruun Mathiasen, and Pernille Steen Pedersen. "The End of Greenwashing? A Research Agenda for the Study of Legitimate Corporate Environmentalism." Accessed January 5, 2024. https://research.cbs.dk/files/96633941/pedersen_et_al_greenwashing_publishersversion.pdf.

Pendse, Meenal Kaustubh, Varsha Shriram Nerlekar, and Pooja Darda. 2022. "A Comprehensive Look at Greenwashing from 1996 to 2021: A Bibliometric Analysis." *Journal of Indian Business Research*. doi:10.1108/jibr-04-2022-0115.

Popular Science. "Title of the Article." Accessed February 20, 2024. https://www.popsci.com/science/human-nature-data-comparison/.

CEU eTD Collection

Princeton University. "WordNet." Accessed February 21, 2024. https://wordnet.princeton.edu/.

Qader, Wisam, Musa M. Ameen, and Bilal Ahmed. 2019. "An Overview of Bag of Words; Importance, Implementation, Applications, and Challenges." In *Proceedings of the 2019 International Electronics Communication Conference*, 200-204. https://doi.org/10.1109/IEC47844.2019.8950616.

Raymond, Eric S. 1999. *The Cathedral and the Bazaar*. Sebastopol, CA: O'Reilly Media.

Reimers, Nils, and Iryna Gurevych. 2019. "Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks." *arXiv* preprint arXiv:1908.10084. https://arxiv.org/abs/1908.10084.

Santos, C., A. Coelho, and A. Marques. 2023. "A Systematic Literature Review on Greenwashing and Its Relationship to Stakeholders: State of Art and Future Research Agenda." *Management Review Quarterly*. https://doi.org/10.1007/s11301-023-00337-5.

ScienceDaily. "Title of the Article." Accessed December 5, 2023. https://www.sciencedaily.com/releases/2013/05/130522085217.htm.

ScienceDirect. "Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence." Accessed May 9, 2024. https://www.sciencedirect.com/science/article/pii/S277266222300070X.

Scientific American. "American Consumption Habits." Accessed May 9, 2024. https://www.scientificamerican.com/article/american-consumption-habits/.

Smith, Katherine, and Tracy Brower. 2012. "Longitudinal Study of Green Marketing Strategies That Influence Millennials." *Journal of Strategic Marketing* 20. https://doi.org/10.1080/0965254X.2012.711345.

Soergel, B., E. Kriegler, I. Weindl, et al. 2021. "A Sustainable Development Pathway for Climate Action within the UN 2030 Agenda." *Nature Climate Change* 11: 656–664. https://doi.org/10.1038/s41558-021-01098-3.

Stammbach, Dominik, Nicolas Webersinke, Julia Bingler, Mathias Kraus, and Markus Leippold. 2022. "Environmental Claim Detection." SSRN. https://ssrn.com/abstract=4207369 or https://doi.org/10.2139/ssrn.4207369.

Strand, R. 1983. "A Systems Paradigm of Organizational Adaptations to the Social Environment." *Academy of Management Review* 8 (1): 90–96.

Strycharz, Joanna, Nadine Strauss, and Damian Trilling. 2018. "The Role of Media Coverage in Explaining Stock Market Fluctuations: Insights for Strategic Financial

CEU eTD Collection

Communication." *International Journal of Strategic Communication* 12 (1): 67-85. https://doi.org/10.1080/1553118X.2017.1378220.

Sun, Xiaofei, Yuxian Meng, Xiang Ao, Fei Wu, Tianwei Zhang, Jiwei Li, and Chun Fan. 2022. "Sentence Similarity Based on Contexts." *Transactions of the Association for Computational Linguistics* 10: 573–588.

Sun, Yongbo, and Binbin Shi. "Impact of Greenwashing Perception on Consumers' Green Purchasing Intentions: A Moderated Mediation Model." *Sustainability*, Business School, Beijing Technology and Business University, Beijing 100048, China. Correspondence: sb990826@163.com.

Talpur, S., M. Nadeem, and H. Roberts. 2023. "Corporate Social Responsibility Decoupling: A Systematic Literature Review and Future Research Agenda." *Journal of Applied Accounting Research*.

Tang, P., C. Wang, Q. Jiang, X. Liu, and J. Wang. 2023. "Symbol or Substance? Environmental Regulations and Corporate Environmental Actions Decoupling." *Journal of Environmental Management* 346: 118950.

TechTarget. "BERT Language Model." Accessed February 21, 2024. https://www.techtarget.com/searchenterpriseai/definition/BERT-language-model.

TechTarget. "Vector Embeddings." Accessed May 9, 2024. https://www.techtarget.com/searchenterpriseai/definition/vector-embeddings.

Tetlock, Paul. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *Journal of Finance* 62: 1139-1168. https://doi.org/10.2139/ssrn.685145.

The Data Warehousing Institute (TDWI). "All Trustworthy Data Requires the Same Three Things." Accessed December 12, 2023. https://tdwi.org/articles/2017/12/08/diq-all-trustworthy-data.aspx.

The Guardian. "Coca-Cola Admits It Produces 3m Tonnes of Plastic Packaging a Year." The Guardian, March 14, 2019. Accessed February 15, 2024. https://www.theguardian.com/business/2019/mar/14/coca-cola-admits-it-produces-3m-tonnes-of-plastic-packaging-a-year.

The Wall Street Journal. "Global Executives Say Greenwashing Remains Rife." Accessed December 15, 2023. https://www.wsj.com/articles/global-executives-say-greenwashing-remains-rife-10a0e273?mod=hp_minor_pos16.

Tiwari, Ashish. 2022. "Artificial Intelligence and Machine Learning for EDGE Computing." In *Analyzing the Leading Causes of Traffic Fatalities Using XGBoost and Grid-Based Analysis: A City Management Perspective*. Available from:

CEU eTD Collection

https://www.researchgate.net/figure/Calculation-of-Precision-Recall-and-Accuracy-in-the-confusion-matrix_fig3_336402347 [accessed May 23, 2024].

Tuzzolino, F., and B. R. Armandi. 1981. "A Need-Hierarchy Framework for Assessing Corporate Social Responsibility." *The Academy of Management Review* 6 (1): 21–28.

Underwriters Laboratories (UL). "The Sins of Greenwashing." Accessed November 30, 2023. https://www.ul.com/insights/sins-greenwashing.

United Nations. "Greenwashing." Accessed February 11, 2024. https://www.un.org/en/climatechange/science/climate-issues/greenwashing.

United Nations. "Net-Zero Coalition." Accessed March 3, 2024. https://www.un.org/en/climatechange/net-zero-coalition.

United Nations. 2023. "Goal 12: Responsible Consumption and Production." United Nations Sustainable Development. Accessed May 1, 2024. https://www.globalgoals.org/goals/12-responsible-consumption-and-production/.

United States Environmental Protection Agency (EPA). "Learn About Volkswagen Violations." Accessed April 5, 2024. https://www.epa.gov/vw/learn-about-volkswagen-violations.

University of Stavanger. "Title of the Document." Accessed March 5, 2024. https://uis.brage.unit.no/uis-xmlui/handle/11250/3090545.

Urbański, Mariusz, and Adnan ul Haque. 2020. "Are You Environmentally Conscious Enough to Differentiate between Greenwashed and Sustainable Items? A Global Consumers Perspective." *Sustainability* 12 (5): 1786. https://doi.org/10.3390/su12051786.

Varini, F. S., Jordan Boyd-Graber, Massimiliano Ciaramita, and Markus Leippold. 2020. "ClimaText: A Dataset for Climate Change Topic Detection." *Tackling Climate Change with Machine Learning (Climate Change AI) Workshop at NeurIPS*, December 2020.

Visser, Wayne, and Jeffrey Hollender. *The Age of Responsibility: CSR 2.0 and the New DNA of Business*.

Wartick, S. L., and P. L. Cochran. 1985. "The Evolution of the Corporate Social Performance Model." *Academy of Management Review* 10 (4): 758–769.

Wood, D. J. 1991. "Corporate Social Performance Revisited." *The Academy of Management Review* 16 (4): 691–718.

World Wildlife Fund (WWF). "Guide to Greenwashing." Accessed February 14, 2024. https://www.wwf.org.uk/learn/guide-to-greenwashing.

CEU eTD Collection

Yang, Zhi, Thi Nguyen, Nam Nguyen, Thi Nguyen, and Thi Cao. 2020. "Greenwashing Behaviours: Causes, Taxonomy and Consequences Based on a Systematic Literature Review." *Journal of Business Economics and Management* 21: 1486-1507. https://doi.org/10.3846/jbem.2020.13225.

Zhang, H., and M.O. Shafiq. 2024. "Survey of Transformers and Towards Ensemble Learning Using Transformers for Natural Language Processing." *Journal of Big Data* 11 (1): 25. https://doi.org/10.1186/s40537-023-00842-0.

CEU eTD Collection