Unveiling Green Facades: Detecting Greenwashing Tendencies in Corporate Sustainability Reports

Marvin Motz
Karlsruhe Institute of Technology
marvin.motz@kit.edu

Seyyid Uzun
Karlsruhe Institute of Technology
seyyid.uzun@kit.edu

Anuja Hariharan
Karlsruhe Institute of Technology
anuja.hariharan@kit.edu

Christof Weinhardt
Karlsruhe Institute of Technology
weinhardt@kit.edu

Abstract

Greenwashing poses a significant challenge to the fight against climate change by undermining trust in corporate sustainability claims. This study introduces the Greenwashing Tendency Score (GTS), an automated method designed to detect greenwashing tendencies in corporate sustainability reports. By leveraging sentiment and alignment analysis techniques in conjunction with ESG ratings, the GTS quantifies discrepancies between reported and actual sustainability performance. We applied our methodology to 36 DAX40 companies over the years 2020 to 2022. Our key findings reveal substantial variations in greenwashing tendencies among these companies, emphasizing the need for more transparent and reliable sustainability reporting. The GTS emerged as a scalable, reproducible, and objective tool that can aid investors, regulators, NGOs, and companies in identifying greenwashing practices. This research not only contributes to the sustainable finance literature but also provides actionable insights for stakeholders committed to fostering accountability and integrity in corporate sustainability communications.

Keywords: Greenwashing, Detection, NLP, Sustainability, SDG

1. Introduction

The fight against climate change faces several hard to overcome obstacles and is in dire need for large amounts of capital (IPCC, 2022; UNFCCC, 2024). Although states and other public entities worldwide are investing substantial sums into sustainable projects, the public sector alone cannot provide all the necessary investments (Bhandary, 2022; UN, 2023). Thus, the concept of sustainable finance has emerged to direct more capital towards companies and projects that are committed to creating positive impacts on the environment and society, rather than negative ones (Fatemi & Fooladi, 2013). To allocate these funds effectively, sustainable investors must be able to differentiate between projects and businesses that align with their sustainability goals and those that do not. Organizations, however, have various incentives to portray their sustainability performance more positively than it actually is (Delmas & Burbano, 2011; Lyon & Montgomery, 2015), a phenomenon widely known as greenwashing. This results in a typical principal-agency problem between sustainable investors and capital seeking companies (Eisenhardt, 1989; Ghitti et al., 2023; Kimmerle, 2019).

Not least because of the resulting agency costs, greenwashing is a major hurdle in the fight against climate change and has become an increasingly widespread occurrence (Montgomery et al., 2023). In 2021, the European Commission released a report that screened websites of consumer products and found false or deceptive sustainability claims in 42% of cases (EC, 2021). In Europe and the Americas, the number of greenwashing cases in the banking and financial services sector rose by 70% between 2022 and 2023 (RepRisk, 2023). Regulators have also recognized the issue, with the European Parliament recently introducing legislation to combat greenwashing at the product level (EP, 2024).

Although these reports have focused mainly on greenwashing at the product level primarily concerning consumers, investors are also affected by greenwashing of companies. In recent years, it has become standard practice to incorporate Environmental, Social, and Governance (ESG) factors into investment decisions (Amel-Zadeh & Serafeim, 2018; Krueger et al., 2020),

URI: https://hdl.handle.net/10125/108951
978-0-9981331-8-8
(CC BY-NC-ND 4.0)

creating the need for reliable sustainability reporting for this stakeholder group as well. The motives behind this practice range from moral and ethical considerations to purely financial reasons (Krueger et al., 2020). Regardless of these motives, greenwashing can harm investors by violating their trust, damaging their reputation, and also causing severe financial losses. A notable example is the Volkswagen diesel scandal, which resulted in significant financial and reputational costs for the company, and consequently, its shareholders (Barth et al., 2022). Volkswagen's misconduct even adversely affected the stock prices of other car manufacturers, thereby financially harming their shareholders as well (Barth et al., 2022; Trefis, 2015). Greenwashing undermines trust among both consumers and investors, hindering the efficient allocation of capital towards sustainable companies and projects which is essential for the transition towards a more sustainable economy (Gatti et al., 2021; Wang et al., 2020; Yang et al., 2020).

The widespread dissemination of greenwashing and the resulting repercussions necessitate automated detection systems to help investors, consumers and litigators combat this malpractice. However, several issues make greenwashing detection exceedingly difficult. First, there is no universally accepted definition of greenwashing (de Freitas Netto et al., 2020; Nemes et al., 2022). While most definitions agree that greenwashing involves a discrepancy between an entity's actual sustainability performance and its exaggeratedly positive communication, the specific manifestations of greenwashing typically vary across four dimensions. While most scholars restrict greenwashing to **environmental claims** (e.g., Delmas and Burbano (2011); Ramus and Montiel (2005)), others also consider, for example, **social performance** (e.g., Lyon and Maxwell (2011)). Another distinction lies in whether greenwashing requires **intentional** deception (e.g., Ramus and Montiel (2005); Bowen and Aragon-Correa (2014)) or if intention is not a necessary prerequisite (e.g., Delmas and Burbano (2011); Walker and Wan (2012)). Additionally, some definitions specify that greenwashing refers to **companies** (e.g., Delmas and Burbano (2011)) while others apply the term more broadly any organization (e.g., Lyon and Montgomery (2015)). Furthermore, greenwashing can be differentiated based on whether it occurs at the **product level** or the **organizational level** (e.g., Delmas and Burbano (2011); de Freitas Netto et al. (2020)).

A second obstacle to detecting greenwashing lies in the diverse methods used for its perpetration (de Freitas Netto et al., 2020; Lyon & Montgomery, 2015; Siano et al., 2017; TerraChoice, 2010). These methods often require different detection techniques, particularly when considering the differences between text-based and visual-based greenwashing. Finally, and partly due to the aforementioned challenges, there is limited data available on greenwashing (Oppong-Tawiah & Webster, 2023). This scarcity of data hinders the applicability of some of the data-driven approaches, such as supervised machine learning methods, which are commonly used in related fields like fake news detection (Ozbay & Alatas, 2020; Shu et al., 2017).

Consequently, most public exposures of greenwashing still arise from manual investigations, often conducted by journalists (Guardian, 2023) or NGOs (EC, 2024), or are discovered serendipitously when uncovering greenwashing was not the primary target of an investigation (Glinton, 2015). However, these efforts typically expose only the more severe and illegal instances of greenwashing, leaving the subtler and often legally permissible cases under-investigated.

In this study, we develop an automated approach aimed at detecting not only overt instances of greenwashing but potentially also the more subtle but pervasive greenwashing efforts by organizations. For our purposes, we define greenwashing as the discrepancy between an entities (overly positive) sustainability performance communication and its actual sustainability performance. This definition is deliberately broad, encompassing the key aspects of various definitions proposed by other authors. It emphasizes the gap between communicated and actual performance without limiting the type of entity, the sustainability dimension, or assuming intent, as none of these are necessary for our approach. Although this study focuses on companies, our approach should be easily transferable to other entities.

To quantify the components of this greenwashing definition, we employ natural language processing techniques, specifically *sentiment analysis* on companies' sustainability reports and *alignment analysis* between these reports and the UN's Sustainable Development Goals (SDGs), to asses the companies' communications. We then contrast these assessments with ESG ratings from a third-party data provider, which serve as a proxy for actual sustainability performance. This method is applied to 36 companies within the DAX40 stock index over the years 2020, 2021, and 2022.

While we are confident in the measure that we have developed, we deliberately named the resulting indicator the Greenwashing *Tendency* Score (GTS). The primary reason is that these scores are meaningful only in comparison to each other. For instance, a company

with a GTS of 1.5 cannot be interpreted in isolation, it requires context, such as another company having a GTS of 1.0 or the sample mean being 1.2. This implies that we cannot make absolute assertions about the level of greenwashing, as the overall level in the sample could be very low or very high. Highlighting this property of the GTS is particularly important given that even companies with strong sustainability performance may hesitate to report their accomplishments for fear of unfoundedly being labeled as greenwashers (Gatti et al., 2021; Lyon & Montgomery, 2015). Therefore, we emphasize that we are not claiming any of the companies mentioned are necessarily engaged in greenwashing, but rather indicating tendencies that emerge using our approach. Although we could have avoided this issue by anonymizing the results, we decided against it to enhance the much-needed reproducibility and comparability in the greenwashing literature for other researchers.

The rest of the paper begins with an overview of existing greenwashing detection approaches and related literature. Next, we provide a detailed explanation of our methodology and data sources. Our results are presented in Section 4 and discussed in Section 5. Finally, we summarize our findings, highlight the study's limitations and suggest future research avenues in the last section.

2. Related Work

While there is considerable literature on the concept of greenwashing and its various forms (e.g., Lyon and Montgomery (2015); de Freitas Netto et al. (2020); Gatti et al. (2021); Seele and Gatti (2017)) the (automated) greenwashing detection literature is still in its infancy (Moodaley & Telukdarie, 2023). One of the earlier works in this field is by Aggarwal and Kadyan (2014), where the authors manually scored 40 companies based on different greenwashing criteria in their advertisements, websites or sustainability reports, and then compared these scores to their Corporate Social Responsibility score. Although the authors followed a systematic approach, the manual method has inherent disadvantages introduced by the human element, i.e., subjectivity involved in the scoring and assessment of possible greenwashing claims (Aggarwal & Kadyan, 2014), and the potential for errors that may result in missing certain claims in the companies' communication (Zhou et al., 2004). Finally, the manual approach is time-consuming and labor-intensive, which limits its scalability. More recently, Oppong-Tawiah and Webster (2023) developed a promising method of detecting greenwashing via linguistic cues, inspired by the fake news detection literature. They also show that there is a relation between greenwashing and financial performance, further emphasizing the importance of uncovering organizations' greenwashing tendencies. Their method has the advantage of not requiring a ground truth or extensive manual work, such as annotations, thereby significantly reducing potential subjectivity and scaling obstacles. Oppong-Tawiah and Webster (2023) use linguistic cues in Twitter posts, but they also encourage future research to, inter alia, investigate content based methods and other media. In this regard, we follow their recommendation by utilizing content based alignment scores and the more formal medium of sustainability reports.

There are a few more approaches to automated greenwashing detection. Vinella et al. (2023) create a binary classifier using self specified greenwashing characteristics and a variation of the climateBERT model (Webersinke et al., 2021). A Swiss company developed an approach that compares the sentiment of forward and backward-looking news reports about a company, accrediting a high greenwashing risk when the forward-looking sentiment is higher compared to the backward-looking sentiment (Mach & Rochat, 2022). Finally, there was a related shared task at the SwissText 2023 conference that was also aimed at greenwashing detection. They compare sentiment analysis of internal documents (i.e., sustainability reports) with the sentiments of external documents (e.g., media reports) in a first step, while repeating the same with SDG alignment in a subsequent step (Contessotto et al., 2023).

3. Method and Data

We developed our methodology based on our before mentioned definition of greenwashing that builds on the intersection of already existing concepts of greenwashing in literature. To quantify the resulting components of greenwashing, we thus need to find variables that adequately capture the way that companies communicate their sustainability efforts and their real sustainability performance.

For the latter we use ESG ratings from a third party data provider (LSEG, formerly Refinitiv). While these kind of ratings have their limitations (see e.g., Berg et al. (2020, 2022), they are still considered one of the better proxies for sustainability performance and are commonly used, for example, in the sustainable finance literature (e.g., Coelho et al. (2023) and Pástor et al. (2022). Although, external media outlets might be a viable alternative and are used in related studies (Contessotto et al., 2023; Mach & Rochat, 2022), their coverage often exhibits biases, such as favoring

advertisers (Beattie et al., 2021; Blasco & Sobbrio, 2012; Ellman & Germano, 2009), larger companies (Fang & Peress, 2009; Nokelainen & Kanniainen, 2018), or salient industries during crises (Jonkman et al., 2018).

For the assessment of the organizations communication, we use two different data sources and indicators and our sample consists of the DAX40 companies for the years of 2020, 2021 and 2022. First, we use the companies' sustainability reports that we gathered from their websites, in a few cases where this was not possible we requested them via email or extracted corresponding sustainability related parts from annual reports. For some companies (Daimler Truck, Porsche AG, Porsche Automobil Holding SE and Siemens Healthineers) we were unable to obtain reports for every year, thus, we excluded them from our analysis. Sustainability reports are likely the most formal and comprehensive communication of a company's overall sustainability efforts on an organizational level that is currently publicly available. In this regard, they differ from more informal and targeted forms of communication such as social media posts or advertisements and are thus more suitable for assessing a companies overall greenwashing tendencies.

The second source is the description of the SDGs (UN, 2015). The SDGs are widely accepted international guidelines that cover a broad range of sustainability issues, including environmental, social, and governmental goals, and are often used as a benchmark (Elalfy et al., 2020; Pattberg & Widerberg, 2016; Stevens & Kanie, 2016). Additionally, investors who aim to make a positive impact also utilize the SDGs to guide their sustainable investments (Bauer et al., 2021; Paetzold et al., 2022), demonstrating their relevance for this particular stakeholder group. Therefore, we use the SDGs as an objective and comprehensive framework for sustainability to facilitate our content comparisons.

Our first communications indicator, similar to some related studies (Contessotto et al., 2023; Mach & Rochat, 2022; Vinella et al., 2023), is sentiment analysis to capture the tone and emphasis of the companies' sustainability reports. To increase the reliability of the results we first apply multiple text preprocessing methods. We start by normalizing the texts as much as possible. They are transformed to all lowercase and we removed diacritics, extra white spaces, URLs, contact information, named entities and contact details, such as email addresses and phone numbers. Furthermore, we expanded abbreviations and contractions. We then use the Natural Language Toolkit (Bird et al., 2009) to perform tokenization, lemmatization, part-of-speech tagging and removal of stop words. After these preprocessing steps we apply the widely recognized VADER (Hutto & Gilbert, 2014) sentiment analysis model on a sentence level. While VADER was initially developed for social media text, its performance was also affirmed for more formal and longer texts such as news outlets (Castellanos et al., 2021) and is therefore also suitable for our purposes.

However, sentiment analysis alone only captures the emotional tone of a text and does not assess the relevancy of the content. Thus, we apply a second measure to account for the content-related aspects of the companies' communications. Without this additional measure, it would be possible for two companies to receive the same greenwashing score, even if one company boasts about every aspect of sustainability (high sentiment) but delivers on only a few aspects (lower ESG score), while another company communicates truthfully about the sustainability aspects where they actually have achievements (also high sentiment and lower ESG score). Intuitively, most people would likely view the first case as exhibiting more greenwashing.

Our solution involves a second factor in the communication assessment: an alignment score that compares the content of the sustainability reports with that of the description texts of the SDGs. We employ the pretrained word embedding algorithm *flax-sentence-embeddings/all_datasets_v3_mpnet-base* from Hugging Face that is based on the MPNet language model (Song et al., 2020) on a sentence level. After aggregating the sentence vectors to single document vectors their distance to the corresponding SDG vector is then calculated via the popular cosine similarity measure (e.g., Gunawan et al. (2018), Lahitani et al. (2016), and Sitikhu et al. (2019)).

These three measures - sentiment, alignment and ESG score - form the basis of our Greenwashing Tendency Score (GTS) that is calculated for each company and year. The GTS can be summarized by the following formula:

$$GTS_i = \frac{SV_i * SDGA_i}{(\frac{ESGS_i}{100})^2} * 10$$

With the index i representing the respective year for which the report and the ESG score are obtained, SV being the computed sentiment value, $SDGA$ the corresponding alignment score (cosine similarity) between the sustainability report and the SDGs, and $ESGS$ the ESG score provided by our data provider LSEG. By design, SV and SDGA can both take on values between -1 and 1 indicating a negative or

Table 1. Summary statistics for the single factors of the GTS from 2020 to 2022

	SV 2020	SV 2021	SV 2022	SDGA 2020	SDGA 2021	SDGA 2022	ESGS 2020	ESGS 2021	ESGS 2022
mean	0.25	0.24	0.23	0.37	0.38	0.37	82.33	83.26	82.86
std	0.04	0.04	0.04	0.05	0.05	0.04	8.3	7.55	7.47
min	0.17	0.17	0.16	0.18	0.22	0.26	62.54	61.78	63.78
max	0.37	0.4	0.33	0.44	0.47	0.46	94.31	94.41	93.44

positive sentiment and very low or very high alignment, respectively. However, we expect the sentiment score to be non-negative, since companies are unlikely to use negative language when talking about their own sustainability efforts. Similarly, given the prescribed common topic of sustainability we also expect the SDGA to be non-negative.

The scale of the ESGS, on the other hand, is prescribed by the data provider and ranges from 0 - 100. Therefore, we divide these scores by 100 to align them with the scale to the numerator in the fraction. Additionally, we square the denominator to match multiplication of the two factors in the nominator. This ensures that both communication and actual performance are equally represented in the GTS. Finally, to achieve a more convenient overall scale, we multiply by 10. As we will show in the next section, this results in GTS values that are approximately between 0.7 and 3, with higher values indicating a higher greenwashing tendency.

4. Results

To calculate the GTS we first compute the SV, SDGA and ESGS scores. The summary statistics for these scores are shown in Table 1. As expected, both SV and SDGA have positive values, ranging from 0.17 to 0.4 and 0.18 to 0.47, respectively. Their means and standard deviations remain very stable over the three years, indicating no clear trends during this period. The same stability applies to the ESGS, which ranges between 62.54 and 94.41, with its means and standard deviations also remaining consistent over the considered period.

To understand the sentiment scores assigned to different sentences, Table 2 shows randomly drawn sentences from the 10% of most positive, negative and neutral sentences from the year 2020. The positive and neutral sentences appear to be correctly assessed. Given that sustainability reports often include many tables and figures, it is likely that many of the neutral sentences are short snippets, possibly part of a table or figure labeling, such as the one shown in this case. The negative example, however, appears counterintuitive since the company reports on improving their systems. Nevertheless, the prevalence of negative words such as 'fraud' and 'risk' makes it understandable that the algorithm assigned a negative sentiment. This is not necessarily detrimental to the company, as a more negative sentiment value will decrease the GTS. In this particular case, it could also be argued that the need to improve the 'Fraud Risk Management Framework' implies the company acknowledges previous issues, making a negative assessment justified. However, we cannot ensure that the same logic applies to other examples.

Table 2. Example sentences for positive, negative and neutral sentiments

Sentiment	Example
0.765	The discourse on important questions relating to social, economic and financial policy is of great importance to us.
0.0	Group abs
-0.827	In addition, we improved the Fraud Risk Management Framework to reinforce awareness of fraud risks across the business divisions.

Combining the components via the GTS formula gives us the ranking shown in Table 4, with the highest value indicating the highest greenwashing tendency. For conciseness we only show the top and bottom five companies here. Table 3 shows descriptive statistics on the percentage changes of the GTS and its components between 2020 and 2022. On average there is a reduction of -3.64% in GTS from 2020 to 2022, indicating a slight decrease in greenwashing tendency for the DAX40 companies over the three years. The standard deviation at 26.1%, however, shows considerable variation between the companies.

For instance, Beiersdorf showing the largest decrease in greenwashing tendency (-27.13%) despite still having the highest GTS overall. This reduction is

Table 3. Top and bottom five companies by GTS of DAX40 companies from 2020 to 2022

	Company	GTS 2020	GTS 2021	GTS 2022	GTS Avg
1	Beiersdorf	2.79	3.01	1.66	2.49
2	Siemens Energy	2.01	1.92	2.01	1.98
3	RWE	1.96	1.55	2.16	1.89
4	Henkel	1.91	1.74	1.76	1.8
5	Qiagen	1.48	1.5	1.89	1.62
	...				
32	Hannover Rück	1.26	0.89	0.9	1.02
33	Commerzbank	0.85	1.07	1.04	0.99
34	Mercedes-Benz	1.14	0.84	0.91	0.96
35	BMW	1.03	0.8	0.84	0.89
36	Volkswagen	0.84	0.93	0.78	0.85

Table 4. Percentage changes of average GTS and its components between 2020 and 2022

	Δ SV	Δ SDGA	Δ ESGS	Δ GTS
mean	-5.32	2.62	0.9	-3.64
std	11.77	21.78	5.4	26.11
min	-28.78	-27.13	-13.71	-40.49
max	24.78	101.43	13.45	86.36

achieved by lowering their SV by 22.8%, their SDGA by 11.3% and an increase in ESG over the same period of 7.27%, indicating a uniform development towards a lower greenwashing tendency. Conversly, Brenntag, which ranks 8th overall in the GTS ranking, increased their GTS from 1.02 in 2020 to 1.9 in 2022, placing them with the third highest greenwashing tendency in that year. In their case, the change mainly stems from a steep increase of over 101% in SDGA while SV (-5.74%) and ESGS (0.94%) changed only slightly in comparison.

Similarly, Table 5 demonstrates how the GTS captures the different aspects of greenwashing using data from three example companies in 2022. First, Qiagen, despite having an ESGS slightly above the sample mean, ranks 5th over the three years and even 3rd in 2022. This is due to their very positively framed communications, reflected in their SV and SDGA, which are both the second highest for 2022. In contrast, Hannover Rück is among the companies with the lowest greenwashing tendency, despite their relatively low ESGS, primarily due to their more neutral reporting. Lastly, Bayer, which has the highest ESGS in 2022, only ranks in the lower midfield in terms of GTS because their communication factors also have relatively high values.

Table 5. Composition of GTS using the example of three companies in 2022

	Company	SV	SDGA	ESGS	GTS
5	Qiagen	0.3	0.45	84.74	1.89
21	Bayer	0.26	0.39	93.44	1.15
32	Hannover Rück	0.16	0.33	78.14	0.9

5. Discussion

In this study, we developed a novel automated approach to detecting greenwashing tendencies in corporate sustainability reports. Following a general definition of greenwashing, the Greenwashing Tendency Score (GTS) leverages the natural language processing techniques of sentiment and alignment analysis to contrast corporate sustainability communications with actual sustainability performance, as measured by ESG ratings.

The results show that the GTS effectively captures various aspects of greenwashing, i.e., the emotional tone of sustainability communication, its relevance in terms of content and actual sustainability performance. For example, a company with a relatively weak actual sustainability performance can receive a lower GTS than a company with a comparatively strong performance, if it reports more neutrally about their efforts. However, this method also allows for a company that has the highest actual sustainability performance (measured by the ESG score) to only receive a medium rank because their reporting was also more positive than that of the other companies in the sample. It is reasonable in this situation to argue that given their surpassing actual performance, they should also be allowed to report positively about their accomplishments.

While this critique is valid, we believe the GTS still holds significant merit. Firstly, it is a fully automated approach for detecting greenwashing tendencies, which can potentially be scaled to a large number of companies and reporting years with minimal manual labor. Additionally, it relies solely on open-source Python packages for text processing, making it easily reproducible. Secondly, the GTS provides a holistic view by analyzing overall communication rather than focusing on isolated statements or advertisements. Thirdly, while the GTS offers an initial overall indication, it is also a transparent measure that allows users to examine its components and investigate the different factors that play into it separately if necessary. Fourthly, it is largely objective, as the algorithms and data used are widely recognized and have been utilized

in various research projects. This approach requires no additional manual annotations or subjective decisions beyond the chosen methods. Finally, although the GTS currently focuses on sustainability in general, it can be further developed. For example, different factors can be weighted according to personal preferences (e.g., to avoid companies with very high ESGS to still receive high GTS), or the focus can be narrowed to specific sustainability issues by using only relevant SDGs and corresponding ESG sub-scores.

Our study contributes to the sustainable finance and greenwashing detection literature by introducing an automated, scalable, and largely objective method for indicating greenwashing tendencies in corporate sustainability communication. By using open-source software and data, we ensure that our approach is reproducible, allowing for validation and adaptation in future research. Researchers can utilize the GTS to explore the interplay between greenwashing and other corporate behaviors, such as corporate social responsibility initiatives and financial performance. Additionally, our approach and results provide a valuable reference for evaluating and comparing other greenwashing detection methods and tools, thereby fostering further innovation in this critical area.

The findings of this study also have substantial implications for practitioners in sustainable finance and the broader fight against climate change. By quantifying the discrepancy between a company's communicated and actual sustainability performance, the GTS indicates differences between the extent to which some companies might engage in greenwashing compared to others. By pointing out these discrepancies, the GTS can help to detect greenwashing practices, thereby improve trust and integrity in the market and ensuring that investments are directed towards companies that truly adhere to sustainable practices.

More specifically, the GTS provides a valuable tool, e.g., for investors, consumers, regulators, NGOs and companies. Investors and consumers can use the GTS to identify and avoid companies that exaggerate their sustainability performance, enabling more informed and ethical investment and purchasing decisions. Regulators can adopt the GTS to monitor long-term greenwashing trends and the impact of their legislation. NGOs and activists can leverage the GTS to identify less obvious greenwashing practices and further investigate uncertain cases, thereby increasing transparency and accountability. Companies can use the GTS to analyze their own and competitors' communications, improving the reliability and tone of their sustainability reporting to better match their actual performance.

6. Conclusion, Limitations and Future Work

This study introduces a novel, automated approach for detecting greenwashing tendencies in corporate sustainability reports through the development of the Greenwashing Tendency Score (GTS). By leveraging sentiment and alignment analysis techniques in combination with ESG ratings, our method quantifies the gap between a company's reported sustainability efforts and its actual performance. The results demonstrate significant differences in greenwashing tendencies among the investigated DAX40 companies, highlighting the need for more transparent and reliable sustainability reporting. Furthermore, by examining the various components of the GTS, the results help identify the drivers behind greenwashing tendencies, which vary between companies.

The GTS presents a promising tool for various stakeholders, including investors, consumers, regulators, and NGOs, to identify and address greenwashing practices. By enhancing the scalability of greenwashing detection, this approach can contribute to greater accountability and integrity in corporate sustainability communications, fostering a more trustworthy sustainable finance ecosystem.

However, the GTS is not without limitations. The model relies on the quality and availability of ESG ratings, which can be subject to biases and inconsistencies. Additionally, while effective, the sentiment and alignment analyses may not capture the full complexity of corporate communications, especially in cases where subtle greenwashing tactics are employed. Further validation and comparison to other methods are necessary to ensure the robustness of our approach.

Future research could focus on enhancing the GTS by incorporating more sophisticated models, such as GPT-4, to improve the accuracy and depth of text analysis. Expanding the scope of the analysis to include other forms of corporate communication, such as social media posts and advertisements, could provide a more comprehensive view of a company's greenwashing practices. Additionally, validating the GTS by increasing the sample size across different industries, geographical regions and longer time periods would help generalize its applicability and effectiveness. As the number of greenwashing detection indicators and methods increases, future work should also compare the results of individual studies to ensure validity and detect potential discrepancies.

References

Aggarwal, P., & Kadyan, A. (2014). Green washing: The darker side of csr. *International Journal of Innovative Research and Pracices*, *2*, 22–35.

Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use esg information: Evidence from a global survey. *Financial Analysts Journal*, *74*, 87–103.

Barth, F., Eckert, C., Gatzert, N., & Scholz, H. (2022). Spillover effects from the volkswagen emissions scandal: An analysis of stock and corporate bond markets. *Schmalenbach Journal of Business Research*, *74*, 37–76.

Bauer, R., Ruof, T., & Smeets, P. (2021). Get real! individuals prefer more sustainable investments. *Review of Financial Studies*, *34*, 3976–4043.

Beattie, G., Durante, R., Knight, B., & Sen, A. (2021). Advertising spending and media bias: Evidence from news coverage of car safety recalls. *Management Science*, *67*, 698–719.

Berg, F., Fabisik, K., Sautner, Z., Albuquerque, R., Berg, T., Fahlenbrach, R., Koskinen, Y., Liang, H., Malloy, C., Moreau, Q., Rockinger, M., Seru, A., Schnitzler, J., & Zhang, C. (2020). *Is history repeating itself? the (un)predictable past of esg ratings*.

Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of esg ratings. *Review of Finance*, *26*, 1315–1344.

Bhandary, R. R. (2022). National climate funds: A new dataset on national financing vehicles for climate change. *Climate Policy*, *22*, 401–410.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with python*. O'Reilly Media Inc.

Blasco, A., & Sobbrio, F. (2012). Competition and commercial media bias. *Telecommunications Policy*, *36*, 434–447.

Bowen, F., & Aragon-Correa, J. A. (2014). Greenwashing in corporate environmentalism research and practice: The importance of what we say and do.

Castellanos, E., Xie, H., & Brenner, P. (2021). Global news sentiment analysis. *Springer Proceedings in Complexity*, 121–139.

Coelho, R., Jayantilal, S., & Ferreira, J. J. (2023). The impact of social responsibility on corporate financial performance: A systematic literature review. *Corporate Social Responsibility and Environmental Management*, *30*, 1535–1560.

Contessotto, A., Giger, T., & Reichmuth, L. (2023). Detecting greenwashing signals through computational language technologies. https://github.com/syX113/hslu-nlp,

de Freitas Netto, S. V., Sobral, M. F. F., Ribeiro, A. R. B., & da Luz Soares, G. R. (2020, December). Concepts and forms of greenwashing: A systematic review.

Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. *California Management Review*, *54*, 64–87.

EC. (2021). European commission-press release screening of websites for 'greenwashing': Half of green claims lack evidence. Retrieved April 9, 2024, from https://ec.europa.eu/commission/presscorner/detail/en/ip_21_269

EC. (2024). European commission-press release commission and national consumer protection authorities starts action against 20 airlines for misleading greenwashing practices.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. https://www.jstor.org/stable/258191

Elalfy, A., Weber, O., & Geobey, S. (2020). The sustainable development goals (sdgs): A rising tide lifts all boats? global reporting implications in a post sdgs world. *Journal of Applied Accounting Research*, *22*, 557–575.

Ellman, M., & Germano, F. (2009). What do papers sell? a model of advertising and media bias. *The Economic Journal*, *119*, 680–704. https://academic.oup.com/ej/article/119/537/680/5089686

EP. (2024). Stopping greenwashing: How the eu regulates green claims. Retrieved April 9, 2024, from https://www.europarl.europa.eu/topics/en/article/20240111STO16722/stopping-greenwashing-how-the-eu-regulates-green-claims

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *Journal of Finance*, *64*, 2023–2052.

Fatemi, A. M., & Fooladi, I. J. (2013). Sustainable finance: A new paradigm. *Global Finance Journal*, *24*, 101–113.

Gatti, L., Pizzetti, M., & Seele, P. (2021). Green lies and their effect on intention to invest. *Journal of Business Research*, *127*, 228–240.

Ghitti, M., Gianfrate, G., & Palma, L. (2023). The agency of greenwashing. *Journal of Management and Governance*.

Glinton, S. (2015). How a little lab in west virginia caught volkswagen's big cheat. *npr*. https://

www . npr . org / 2015 / 09 / 24 / 443053672 / how - a - little - lab - in - west - virginia - caught - volkswagens-big-cheat

Guardian. (2023). Revealed: More than 90% of rainforest carbon offsets by biggest certifier are worthless, analysis shows. https : / / www . theguardian . com / environment / 2023 / jan / 18 / revealed - forest - carbon - offsets - biggest - provider-worthless-verra-aoe

Gunawan, D., Sembiring, C. A., & Budiman, M. A. (2018). The implementation of cosine similarity to calculate text relevance between two documents. *Journal of Physics: Conference Series*, *978*.

Hutto, C. J., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, 216–225. http : / / sentic.net/

IPCC. (2022). Sixth assessment report - mitigation of climate change. https://www.ipcc.ch/report/ar6/wg3/downloads/report/IPCC_AR6_WGIII_FullReport.pdf

Jonkman, J. G., Trilling, D., Verhoeven, P., & Vliegenthart, R. (2018). More or less diverse: An assessment of the effect of attention to media salient company types on media agenda diversity in dutch newspaper coverage between 2007 and 2013. *Journalism*, *19*, 1608–1627.

Kimmerle, H. (2019). The principal-agent problem within sustainable investing. *New Challenges of Economic and Business Development*, 400–409.

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *Review of Financial Studies*, *33*, 1067–1111.

Lahitani, A. R., Permanasari, A. E., & Setiawan, N. A. (2016). Cosine similarity to determine similarity measure: Study case in online essay assessment. *4th International Conference on Cyber and IT Service Management*.

Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate environmental disclosure under threat of audit. *Journal of Economics and Management Strategy*, *20*, 3–41.

Lyon, T. P., & Montgomery, A. W. (2015). The means and end of greenwash. *Organization and Environment*, *28*, 223–249.

Mach, A., & Rochat, M. (2022, August). Greenwashing risk indicator. https : / / www . covalence . ch / docs / Covalence_GreenwashingRiskIndicator_WhitePaper.pdf

Montgomery, A. W., Lyon, T. P., & Barg, J. (2023). No end in sight? a greenwash review and research agenda.

Moodaley, W., & Telukdarie, A. (2023, January). Greenwashing, sustainability reporting, and artificial intelligence: A systematic literature review.

Nemes, N., Scanlan, S. J., Smith, P., Smith, T., Aronczyk, M., Hill, S., Lewis, S. L., Montgomery, A. W., Tubiello, F. N., & Stabinsky, D. (2022). An integrated framework to assess greenwashing. *Sustainability (Switzerland)*, *14*.

Nokelainen, T., & Kanniainen, J. (2018). Coverage bias in business news: Evidence and methodological implications. *Management Research Review*, *41*, 487–503.

Oppong-Tawiah, D., & Webster, J. (2023). Corporate sustainability communication as 'fake news': Firms' greenwashing on twitter. *Sustainability (Switzerland)*, *15*.

Ozbay, F. A., & Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and its Applications*, *540*.

Paetzold, F., Busch, T., Utz, S., & Kellers, A. (2022). Between impact and returns: Private investors and the sustainable development goals. *Business Strategy and the Environment*, *31*, 3182–3197.

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2022). Dissecting green returns. *Journal of Financial Economics*, *146*, 403–424.

Pattberg, P., & Widerberg, O. (2016). Transnational multistakeholder partnerships for sustainable development: Conditions for success. *Ambio*, *45*, 42–51.

Ramus, C. A., & Montiel, I. (2005, December). When are corporate environmental policies a form of greenwashing?

RepRisk. (2023). On the rise: Navigating the wave of greenwashing and social washing. Retrieved November 2, 2023, from https://www.reprisk.com / news - research / reports / on - the - rise - navigating - the - wave - of - greenwashing - and - social-washing

Seele, P., & Gatti, L. (2017). Greenwashing revisited: In search of a typology and accusation-based definition incorporating legitimacy strategies.

Business Strategy and the Environment, *26*, 239–252.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, *19*, 22–36. http://www.journalism.org/2016/05/26/news-use-across-

Siano, A., Vollero, A., Conte, F., & Amabile, S. (2017). "more than words": Expanding the taxonomy of greenwashing after the volkswagen scandal. *Journal of Business Research*, *71*, 27–37.

Sitikhu, P., Pahi, K., Thapa, P., & Shakya, S. (2019). A comparison of semantic similarity methods for maximum human interpretability. *Artificial Intelligence for Transforming Business and Society*.

Song, K., Tan, X., Qin, T., Lu, J., & Liu, T.-Y. (2020). Mpnet: Masked and permuted pre-training for language understanding. *34th Conference on Neural Information Processing Systems*. https://github.com/microsoft/MPNet.

Stevens, C., & Kanie, N. (2016, June). The transformative potential of the sustainable development goals (sdgs).

TerraChoice. (2010). The sins of greenwashing: Home and familiy edition. https://www.twosides.info/wp-content/uploads/2018/05/Terrachoice_The_Sins_of_Greenwashing_-_Home_and_Family_Edition_2010.pdf

Trefis. (2015). The domino effect of volkswagen's emissions scandal. Retrieved April 9, 2024, from https://www.forbes.com/sites/greatspeculations/2015/09/28/the-domino-effect-of-volkswagens-emissions-scandal/?sh=56b804e0d282

UN. (2015). Transforming our world: The 2030 agenda for sustainable development - a/res/70/1.

UN. (2023). Seeking synergy solutions a new financial system to enable both climate and sdg action. http://creativecommons.org/licenses/bync/4.0

UNFCCC. (2024). Ad hoc work programme on the new collective quantified goal on climate finance workplan for 2024. https://unfccc.int/event/2022-high-level-ministerial-dialogue-on-the-new-collective-

Vinella, A., Capetz, M., Pattichis, R., Chance, C., & Ghosh, R. (2023). Leveraging language models to detect greenwashing. http://arxiv.org/abs/2311.01469

Walker, K., & Wan, F. (2012). The harm of symbolic actions and green-washing: Corporate actions and communications on environmental performance and their financial implications. *Journal of Business Ethics*, *109*, 227–242.

Wang, H., Ma, B., & Bai, R. (2020). The spillover effect of greenwashing behaviours: An experimental approach. *Marketing Intelligence and Planning*, *38*, 283–295.

Webersinke, N., Kraus, M., Bingler, J. A., & Leippold, M. (2021). Climatebert: A pretrained language model for climate-related text. http://arxiv.org/abs/2110.12010

Yang, Z., Nguyen, T. T. H., Nguyen, H. N., Nguyen, T. T. N., & Cao, T. T. (2020). Greenwashing behaviours: Causes, taxonomy and consequences based on a systematic literature review. *Journal of Business Economics and Management*, *21*, 1486–1507.

Zhou, L., Burgoon, J. K., Twitchell, D. P., Qin, T., & Nunamaker, J. F. (2004). A comparison of classification methods for predicting deception in computer-mediated communication.