

ChatGPT discovery of green image damaging information for large production companies

Thimm, Heiko; Rasmussen, Karsten Boye

Published in:
Journal of Cleaner Production

DOI:
[10.1016/j.jclepro.2024.143978](https://doi.org/10.1016/j.jclepro.2024.143978)

Publication date:
2024

Document version:
Final published version

Document license:
CC BY-NC-ND

Citation for pulished version (APA):
Thimm, H., & Rasmussen, K. B. (2024). ChatGPT discovery of green image damaging information for large production companies. *Journal of Cleaner Production*, 478, Article 143978.
<https://doi.org/10.1016/j.jclepro.2024.143978>

Go to publication entry in University of Southern Denmark's Research Portal

Terms of use

This work is brought to you by the University of Southern Denmark.
Unless otherwise specified it has been shared according to the terms for self-archiving.
If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.
Please direct all enquiries to puresupport@bib.sdu.dk



ChatGPT discovery of green image damaging information for large production companies

Heiko Thimm^{a,*}, Karsten Boye Rasmussen^b

^a Pforzheim University, School of Engineering, Tiefenbronnerstr. 65, 75175, Pforzheim, Germany

^b University of Southern Denmark, Department Marketing and Management, Campusvej 55, DK-5230, Odense M, Denmark

ARTICLE INFO

Handling Editor: Zhifu Mi

Keywords:

ChatGPT information discovery
Green image damaging information
ESG score
Global production industry
Explorative research study

ABSTRACT

Language models, particularly transformer-based architectures like ChatGPT, have gained significant attention due to their ability to comprehend and generate human-like text. These capabilities are leveraged to retrieve Green Image Damaging (GID) information about a randomized sample of about 400 of the largest production companies. For each sample company prompts to ChatGPT are used to discover and retrieve information of five company topics: CO2 compensation, greenwashing, environmental scandals, noncompliance with environmental legislation and standards, and legal actions related to environmental violations. Through corresponding data analysis, the study explores differences in extents of GID information for regions and industry sectors using the NACE classification scheme. Based on the extent of obtained information the sample is divided into companies where GID information is discovered and companies without GID information. The two groups are compared in terms of company size and ESG scores. Among other results the data analysis suggests that companies with GID information are larger and have significantly better ESG scores than the companies without GID information.

1. Introduction

Today, for companies it is of key economic importance to be viewed by the greater public as a company that respects the environment and pursues the business goals according to the relevant Sustainable Development Goals (SDG) of the United Nations. Therefore, in the business world substantial annual profit shares are spent on building and maintaining a corresponding environmental sensitive reputation that is often referred in short as a ‘green image’ (Alam and Islam, 2021; Heikkurinen, 2010; Mukonza and Swarts, 2020; Zameer et al., 2024). For these efforts information plays a major role due to its potential power to build but also to damage a green company image (Hossain et al., 2023; Sajid et al., 2024; Zou et al., 2015). Damaging a green image may for example have the result that the greenness (or cleanness) of the company is questioned by the public, that the company is viewed as an ‘environmental polluter’, and that public trust in the company is lost. Like for image damages in general, these reputation effects, among others, can cause financial losses due to changing buying behavior of consumers (Chortareas et al., 2024; Hengboriboon et al., 2022; Sajid et al., 2024; Xiong et al., 2021), lower shareholder values (Wu et al., 2022; Xiao et al., 2022; Xue et al., 2023), and lower ESG scores

(Svanberg et al., 2022).

Many different kinds of negative information related to a company may cause such damages. In this work we refer to such information by ‘green image damaging’ (GID) information and especially investigate GID information about the following issues: greenwashing (Freitas Netto, Sobral, Ribeiro and Da Soares, 2020), CO2 compensation (Fischer and Knuth, 2023), environmental scandals (Ater & yosef, 2020) and environmentally harmful incidents (Caiazza et al., 2023), non-compliance with environmental regulations, and violations of environmental legislation (Xia et al., 2024). Real-world GID information of these categories are discovered through ChatGPT.

As a result of the latest advancements in the field of generative Artificial Intelligence (AI) (Bang et al., 2023; Vaswani et al., 2017), ChatGPT of OpenAI (Y. Liu et al., 2023; Roumeliotis and Tselikas, 2023) with its underlying Large Language Model (LLM) and similar systems now offer unprecedented possibilities to easily discover almost all kinds of company information. That is, everyone with Internet access can obtain environmental sensitive information about business entities. In comparison to search engines, among other advantages, ChatGPT requires less efforts to obtain the information, for example, the user can in natural language describe what information the system should search

* Corresponding author.

E-mail addresses: heiko.thimm@hs-pforzheim.de (H. Thimm), kbr@sam.sdu (K.B. Rasmussen).

for. However, there are a number of issues to be considered by ChatGPT users that are interested in environmental sensitive information (Ray, 2023): 1. a carefully thought and well-formed prompt is required in order to obtain a useful fact-based response which is not invented by the LLM i.e. the result of what is referred by hallucination (Bang et al., 2023; Chelli et al., 2024), 2. the underlying knowledgebase may not be fully up-to-date, 3. the information sources are not explicitly referred in the response. Therefore, users are advised to carefully validate the returned information (Y. Li et al., 2021; Sarvandani, 2023).

Even though the data literacy competency of societies might have improved during recent years (Smolnikova, 2020), it can still be assumed that significant numbers of users are unaware of the above described issues of ChatGPT and other generative AI systems. These users may consider the generated response for building opinions and/or judging companies (Oncioiu et al., 2020). As a result, companies that from an environmental sustainability point of view are doing well may be wrongly viewed as polluters or 'bad companies'. Conversely, companies harming the environment may be wrongly viewed as green companies or clean companies. Furthermore, trust and credibility in ESG scores of ESG rating agencies may deteriorate (Demartini and Pagliei, 2023). These considerations motivated our explorative research study which makes use of a global sample of 500 of the largest production companies. For each of the sample companies ChatGPT was prompted for five categories of GID information. The GID information contained in the ChatGPT response was recorded and fused with further data about the companies such as number of employees, financial numbers, the ESG scores of the S&P rating agency, and also data about the companies' websites and content of various web pages. The resulting analytical data set offers many opportunities to reveal descriptive statistics that are helpful to gain a better understanding of the current extent of existing GID information. The statistics also offer a presentation of how firms for which GID information was discovered - simply referred to in the following as companies with GID information - are distributed over different world regions and different industry sectors. Furthermore, companies with GID information are compared to companies without GID information in terms of the company size, financial power, and ESG scores.

In the next section the existing literature is reviewed. Section 3 outlines the theoretical underpinnings that guide our investigation into GID information. Building upon this foundation, Section 4 offers insights into the significance of GID information within the context of the contemporary environmentally conscious landscape. Section 5 elaborates on the methodology employed, detailing how ChatGPT was utilized to extract GID-related insights across five distinct categories. Subsequently, Section 6 encompasses several subsections: Section 6.1 presents the empirical findings, Section 6.2 undertakes a comparative analysis of companies with and without GID information, and Section 6.3 explores the interplay between multiple categories of GID information. Section 7 not only summarizes the main findings of the study but also outlines future directions for research.

2. Related work

Data about environmental wrongdoing of firms have been collected and analyzed in a number of prior studies. The studies used a variety of different data source and data collection methods. The methods used to obtain insights from the empirical data and to test hypotheses include common statistic methods, text analysis and text mining methods, as well as Machine Learning methods from the field of Artificial Intelligence. Theoretical foundations described by the researchers often refer to communication theories (Bormann, 1980; Ihlen, 2011), institutional theory (Meyer, John W.; Rowan, Brian, 1977), legitimacy theories such as media legitimacy (Aerts and Cormier, 2009), reputation theory (Weigelt and Camerer, 1988), journalism theories (Lueddecke & T. Boykoff, 2016), and attribution theory (Burke, 2022; Weiner, 1986).

This work uses ChatGPT (Y. Liu et al., 2023; Roumeliotis and

Tselikas, 2023) as a single exclusive data source (or 'informant') to discover and collect GID information. However, because ChatGPT leverages data from a non-disclosed number of online data sources, actually the data is most likely discovered in and collected from several sources. Multiple research groups however obtained environmental firm data from a truly single source. For example, in a Chinese study (Wu et al., 2022) the researchers used bulletins of the companies provided by a news network to gain insight into environmental penalties of companies. The data on all environmental incidents were collected, and then a total of 88 companies were selected as the sample. For an investigation of stock market reaction Carpentier and Suret (Carpentier and Suret, 2015) analyzed reports from the front page of the New York Times for half a century. A data set of 161 major environmental and non-environmental accidents was obtained through this approach. The study of Ruban and Yashalova (Ruban and Yashalova, 2022) checked the most 'fresh' and only English versions of the codes of conduct at the official web pages of 100 companies from the Forbes GLOBAL 2000 ranking. This led to data of 82 codes which were investigated to reveal insights about extent and types of pro-environmental behavior prescribed by the companies. Among other results the study found that company location is a ' [...] somewhat important determinant of the prescribed pro-environmental behavior: The West European companies prescribe pro-environmental behavior better than those East Asian [...]'. Also, in this work some outstanding differences between these two regions have been found. Alam and Islam (Alam and Islam, 2021) used interviews for their investigation on green corporate image building and green competitive advantage in the apparel sector of Bangladesh. Out of 340 questionnaires sent to mid and top-level management employees a total of 268 could be used for the empirical analysis that involved 53 enlisted firms.

In the majority of prior explorative research on corporate environmental management and in particular in studies with focus on environmental wrongdoing of firms the data are gathered from several sources. The typical sources in which it is looked for the targeted data include data archives and data bases of governmental agencies, regulators, and NGOs (Burke, 2022; R. Li et al., 2023; Zou et al., 2015), media archives (R. Li et al., 2023; W. Li et al., 2022; Zou et al., 2015), website data of the companies (W. Li et al., 2022; Patten and Crampton, 2003; Portella and Borba, 2020; Thimm and Rasmussen, 2023), and the firms environmental reports (Braam et al., 2016; Kang and Kim, 2022; R. Li et al., 2023; Stacchezzini et al., 2016; Tiscini et al., 2022). ChatGPT offers a possibility to discover and collect GID information for a global sample of companies with relatively less resources. A lot more resources are required to obtain such an empirical data set from several data sources and prior studies are often focused on a particular region (e.g. China (R. Li et al., 2023; Su and Fan, 2021; Zou et al., 2015), Latin America (Portella and Borba, 2020), Europe (Baalouch et al., 2019; Bonson and Escobar, 2002; Heras-Saizarbitoria et al., 2015). Also, some prior research use national samples that are further focused on a particular industry sector such as the above Bangladesh study of Alam and Islam (Alam and Islam, 2021) and a study from three Italian Universities explored environmental practices of the Italian textile, clothing, and leather industry (Resta et al., 2014). Some prior work focuses on one particular company. For the single company environmental information about the company including negative information about incidents is obtained from multiple data sources. A Canadian study (Azizul Islam and Aminul Islam, 2011) examined the environmental disclosure initiatives of a Canada-based multinational oil and gas company following two major environmental blowouts at a gas field. The researchers reviewed news articles about the environmental incidents, the company's communication media, including annual reports, press releases and stand-alone social responsibility report.

This study is intentionally descriptive in nature. It reveals extents of retrieved GID information for different world regions and countries, and different sectors of the global manufacturing industry. Differences among industry sectors with respect to environmental pressure have

been studied by the European Commission (Shailendra and Muehmel, 2010). The report suggests a ranking of the top twenty sectors with the greatest pressure on the environment. Insights into the vulnerability of different sectors as caused by the European Union tax reform are provided by the work of FitzGerald, Keeney, and Scott (FitzGerald et al., 2009). In our study too, empirical investigations of corporate ESG scores are performed. The impact of firm size on the ESG Score has also been investigated in recent research of Drempetic, Klein, and Zwergel (Drempetic et al., 2020). The researchers built an empirical data set of nearly 4,000 companies in the period between 2004 and 2015. Through data analysis with several models they found a significant positive correlation between firm size and ESG scores which is also a finding of this work.

The data analyses of ESG scores and firm size are in this study performed with several subsets of the sample. For this purpose, the sample is divided into companies for which GID information was found and companies for which no GID information was found. In principle, GID information indicates that the company most likely is involved in environmental pollution. Therefore, the subset of companies with GID information can also be broadly referred as assumed polluters and the subset of companies without GID information as assumed non-polluters. A similar division of the study sample is also found in other research. For example, in their study of the dissemination of environmental certifications in Europe Heras-Saizarbitoria et al. (2015) divide a study sample into high polluting companies and low polluting companies. The study results show a higher density of environmental certification in sectors with the greatest environmental impact as compared to sector with a lower environmental impact. Reilly and Hynan (2014) use the Newsweek's Greenest Company 2012 listing to partition their sample into 'green companies' and 'not green companies'. The study uses a relatively small sample of 15 corporations in four different industries. The sample is used to explore communication about corporate sustainability on social media platforms and in corporate social responsibility reports.

Other empirical studies aim to explore the impact of negative environmental incidents and penalties, respectively, from several perspectives such as increased media attention to the firm (Zou et al., 2015; Zygliopoulos et al., 2012), more extensive self-disclosure behavior after incidents (Islam and Islam, 2011), lower financial performance (Wu et al., 2022), lower firm value (Fan et al., 2020), increased critical public attention to the firm (Burke, 2022; Wu et al., 2022; Zou et al., 2015), lower environmental performance (Su and Fan, 2021), and stimulation and characteristics of user-generated content (Beedasy et al., 2020; M. Liu et al., 2023).

Zou et al. (2015) investigated how environmental violation events (EVE) harm corporate reputation. The researchers built an empirical data set of 352 environmental violation events in China caused by 330 different firms. The sample was retrieved from the websites of national governments, provincial environmental protection departments and the mass media for the period 2006–2010. The results show that the media regard firms with a history of EVE (Zou et al., 2015), p. 836 '[...] to be more culpable for adverse environmental events and consequently suffer more damage to their reputation, while for firms with a favorable environmental record (environmental awards and honors gained) the reputational harm of an EVE is alleviated to some extent'.

The potential of LLMs such as ChatGPT and possible use cases have been investigated for various application domains including plastic surgery (Gupta et al., 2023), additive manufacturing (Badini et al., 2023), and healthcare (Jin and Dobry, 2023). One of the few works which is in a broader sense is somewhat connected to this research that focuses on corporate sustainability management information is the investigation of ChatGPT as an aid for practitioners of Life Cycle Assessments (LCA) and Life Cycle Impact Assessment (LCI) (Cornago et al., 2023). Based on a SWOT analysis the researchers describe some suggestions of potential application for Transformers in LCA. For example, they suggest that through the analysis of past LCI and LCA impacts data the model could identify where and when environmental performance

improvements would have the greatest impact. Furthermore, they describe that Transformers' forecasting capabilities could improve the potential in supporting LCA impacts reduction strategies by forecasting parts of the LCI. In another research journal article on ChatGPT in the domain of environmental science ChatGPT was used to generate 100 important questions for biology and environmental science (Agathokleous et al., 2023).

3. Foundation of research

In most parts of the world environmental damage and climate change are topics with high media coverage. A growing number of people are becoming more and more afraid of and directly affected by this millennium problem. It is often said that companies of the production industry have in the past done and still presently do damage to the environment through various outputs of their business operations (i.e. development and design, production, distribution) (Cai and Yan, 2021; Wang et al., 2018). Typical outputs include wastewater, exhaust gas, CO₂, odors, waste, noise and vibration. The outputs can cause various adverse environmental impacts such as global warming, resource depletion, air pollution, water pollution, impacts of waste, noise and vibration, bad odors, hazardous chemical pollution. In general, in this research we refer to these companies as polluters. The opposite i.e. non-polluting companies are referred to as clean companies whereas clean is used in a broader sense with the meaning of having a significantly lower environmental impact than polluters.

It is important to emphasize in this research the fact that many of the companies of the production industry have in recent years started to transform their products and operations to a significant lower environmental impact (Eppinger et al., 2021; Labucay, 2022; Lu et al., 2023; OECD, 2015). Obviously, this transformation goes usually hand in hand with efforts to build and maintain a green company image and being viewed as a clean company. However, a green company image can easily be destroyed by information that may in a more or less subtle way associate the company with environmental damage. As described above, in this work we refer to this information by GID information and focus in particular on information concerning greenwashing, CO₂ compensation, environmental scandals and environmentally harmful incidents, non-compliance with environmental regulations, and violations of environmental legislation. Typically, GID information is published by the media (Luedecke & T. Boykoff, 2016; Mateu and Domínguez, 2019), environmental activist groups, NGOs, research institutes, and governmental authorities. As evident in some extreme examples of corporate environmental misbehavior in the past such publications may instantly result into a polluter image (Hossain et al., 2023). But, also cover-up reports about environmental wrongdoings decades ago can still severely damage the company image.

A main consideration of this research is that published reports with GID information are available in media archives, online databases of authorities (e.g. databases of the Environmental Protection Agency of the United States) and research institutes (e.g. Political Economy Research Institute of the University of Massachusetts Amherst (PERI, 2023)), and other special websites. Accessing this steadily growing pool of validated and non-validated critical information is to a large degree free of charges. It is noteworthy to point out that this opportunity to obtain GID information about companies has already been available for more than a decade. One can assume that the interest for this information pool has been growing over the years.

So far, GID information has most likely been explored by interested parties through usual Internet search engines (Thimm and Rasmussen, 2021) or specialized search frontends offered by respective online databases. These approaches impose limitations which can make selective information search and retrieval a time-consuming effort. Since recently, ChatGPT (Roumeliotis and Tselikas, 2023) is available to the general public which based on an underlying LLM offers unprecedented possibilities to search and retrieve 'any kind of information' with little efforts

(Y. Liu et al., 2023). Never before has it been as convenient for the general public as it is today to discover GID information. One can through natural language instructions conveniently demand GID information from ChatGPT. Also, the system offers response in natural language text which may directly contain and explain GID information. Compared to search engine results, users do not need to explore web links contained in the hit list. That these hit lists sometimes contain commercial advertisement links can require extra efforts for the exploration of GID information through search engines. These considerations have been the point of departure of this empirical study that builds on a multidimensional analytical data set and that makes use of quantitative methods. The study is intended to collect and explore GID information for a set of sample companies and to identify various groups of companies with discovered GID information. The design of the study method was influenced by the intention to use a data collection process that imitates how ordinary users are likely to use ChatGPT to discover GID information through ChatGPT. Through corresponding analyses of the aggregated data, the study aims to provide answers to the following research questions.

- RQ1: What amount of GID information can be discovered with ChatGPT?
- RQ2: To which extent is the GID information substantiated by clear facts and to which extent is it rather linked to fuzzy result statements?
- RQ3: To which regions/countries and production sectors do companies with a large extent of GID information belong to?
- RQ4: How do companies for which no GID information is found compare to companies for which GID information was found concerning ESG scores and company size?

The study results offer an indication of how ChatGPT currently presents some major facets of environmental awareness of the global production sector to the world. It can be assumed that this characterization of environmental awareness is not only of interest to the research community. In particular it is assumed that the study insights are also of interest for environmental concerned individuals, special industry unions, media representatives, politicians, and also pressure groups. Insights from this characterization may also alert the reputation management teams of companies because for some of the sample companies the promoted green company image and/or their ESG scores contradicts with the ChatGPT response obtained in our study.

The target group of this empirical business research are large and mostly internationally operating companies of the global manufacturing industry. An initial random sample of about 600 companies was extracted from the ORBIS database (Bureau van Dijk, 2019). The NACE Code (EUROSTAT, 2008) was used to identify companies of a set of focused manufacturing sectors described below. The further selection criteria for company size were: number of employees >1100, Assets >1.000 M USD, Turnover >750 M USD.

In addition to general company data also financial numbers of the last 10-year time period, and the address of the corporate website were extracted from ORBIS. Then, the company's respective ESG scores were extracted and added to the initial data set. The S&P ESG scores and the S&P sustainability scores were extracted from the Bloomberg database. Obviously, the data set was also extended by the set of GID data obtained for the sample companies through the observation method described below.

From the initial sample we selected the sectors displayed in Table 1 which contained high numbers of companies and excluded the smaller remaining sectors. A then following exclusion of regions with very low numbers of companies lead to the sample's focus regions and countries contained in Table 2. The frequency distribution of the resulting final sample of 375 companies concerning the focused industry sectors and focused regions can be obtained from Table 3.

Table 1
Focused manufacturing sectors of study and respective frequencies of sample companies.

Sector	Number of Companies	NACE codes (first two digits)	Description of sector
CEO	145	26	computer, electronic and optical products
VTE	92	29–30	motor vehicles, trailers, semi-trailers, other transport equipment
ME	79	28	machinery and equipment not elsewhere classified
BM	74	24	Basic metals
All	390		

Table 2
Focused regions of study and respective frequencies of sample companies.

Region	Region Short	Number of companies	Countries and number of companies
China	C	131	China 131
Europe	E	93	United Kingdom 10, Germany 12, France 10, Italy 5, Spain 1, Netherlands 3, Switzerland 4, Austria 1, Sweden 4, Belgium 2, Cyprus 1, Finland 4, Ireland 2, Luxembourg 3, Russian Federation 9, Poland 1, Hungary 1, Romania 5, Ukraine 3, Virgin Islands (British) 1, Cayman Islands 6, United States of America 69, Canada 13
North America	A	82	Australia 4, India 6, Japan 24, South Korea 9, Taiwan 14, Hong Kong 3, Singapore 3, Indonesia 1, Thailand 1, Malaysia 3, Philippines 1,
Asia-Pacific (APAC) except China	P	69	
All		375	

Table 3
Frequency distribution of focused sample companies for sectors and regions.

Sector	Region				All
	APAC	China	Europe	North America	
BM	14	28	14	12	68
CEO	28	49	31	34	142
ME	6	29	24	20	79
VTE	21	25	24	16	86
All	69	131	93	82	375

4. Reputation damaging information in Today's environmentally aware world

Company efforts to build and maintain a 'green image' obviously suffer setbacks when GID information becomes public (Hossain et al., 2023). In reaction to environmental wrongdoing described in GID information, people might blame the company for polluting and damaging the environment, and contributing to global warming and the climate change crisis. In the contemporary history of the business world numerous examples of GID information are available on various publication channels. A few prominent examples includes the 'Diesel Gate' scandal (Ater & yosef, 2020), the world-wide pollution scandals of the Oil and Gas Industry, and the case of consumer misinformation, that a German brewery has been accused of (Lütge, 2018).

In general, GID information may have many different facets. Often it revolves around site incidents and company behavior causing direct threats for the environment (e.g. leakages of industrial chemicals) and/or having negative impacts on the environment in the long run (e.g. use of critical chemicals in products or production processes). In a pragmatic attempt to conceptualize information categories that model the various

facets, this study divides GID information into five categories.

For the five proposed GID categories contained in Table 4, the category names are intended to identify into which reputational direction a company image will likely be pushed by GID information about the particular wrongdoing addressed by the category (most right column of Table 3). The particular direction can be roughly obtained from the corresponding self-explanatory category name and the further content of the table. Some of the category identifiers correspond to reputation names sometimes used in press articles or reports of interest groups to refer to companies that have been accused of the corresponding wrongdoing described in the table. The order of the categories in the table suggest a naïve assumption of the image damaging potential of the categories. That is, of the five GID categories for the compensators the lowest damaging potential is assumed while the strongest damaging potential is assumed for the violators.

In the remainder of this article the focus is on the sample companies which are the target of the empirical ChatGPT study. For example, this includes an investigation of the proportions of the five different GID categories. Note that for a given company one may find information of several or even all of the categories which means that a company may be viewed as both compensator, greenwasher, and violator.

5. ChatGPT method to discover GID information

In explorative studies of environmental awareness and environmental disclosure of companies (Chiu et al., 2020; Clarkson et al., 2008; Patten and Crampton, 2003; Thimm and Rasmussen, 2021), the targeted information is often obtained from several sources including environmental reports (Stacchezzini et al., 2016; Tiscini et al., 2022), corporate websites (Patten and Crampton, 2003; Portella and Borba, 2020; Thimm and Rasmussen, 2023), media archives (R. Li et al., 2023; Su and Fan, 2021), and special databases such as RepRisk (Burke, 2022). In many older studies the data was collected through eye-inspection by trained human observers (Berthelot et al., 2013; Braam et al., 2016; Patten and Crampton, 2003; Portella and Borba, 2020; Thimm and Rasmussen, 2021). In the more recent studies this method is complemented by the use of digital information discovery and analysis tools such as advanced search engines (Zou et al., 2015), web scraping tools (Sozzi, 2023; Thimm and Rasmussen, 2023), and text analysis tools (Gill et al., 2008; Kang and Kim, 2022; Moodaley and Telukdarie, 2023; Moreno and Caminero, 2022). The approaches that use multiple sources and multiple tools are expected to enable a comprehensive view of companies. Unlike these approaches, a particular data source and data collection tool is of special interest for this study. As argued above this study obtains an

Table 4
The five categories of GID information considered in the study.

no.	GID Category	GID content	Wrongdoing of company in the past or present
1	Compensator	Information about voluntary CO2 compensation	buying CO2 certificates to offset a specific amount of CO2 caused by its own
2	Greenwasher	Information about greenwashing	using communication and marketing measures to build and maintain a green image that is not substantiated by respective activities (green efforts)
3	Scandalous	Information about environmental scandals and/or incidents	environmental wrongdoing largely perceived as a scandal by the general public
4	Noncompliant	Information about non-compliance with environmental legislation and standards	breaching environmental legislation, standards, or self-commitments
5	Violator	Information about legal actions related to environmental violations	environmental wrongdoing leading to legal allegations with possible punishment

indication of extent of information for the above five categories of GID information under the premise that the information is exclusively obtained through ChatGPT.

In a first experimental step a number of tests were performed in order to determine the suitable ChatGPT prompts that yield the targeted GID information. The test results revealed that potential confusion by ChatGPT of company names can be avoided by the specification of both the company name and the web address of the company's web site. Based on the test results the five prompts shown in Table 5 were finalized for the data collection. Obviously, the example company name and url (i.e. 'Yieh Phui Enterprise' and 'www.yiehphui.com.tw') in Table 5 were in the later data collection step repetitively replaced by information of the currently investigated company of the sample.

The data collection was performed by a group of 56 individuals including members of the research project and third semester research assistants of an academic study program. After comprehensive training the data collection was completed according to given instructions during the period 14th March to 15th April May 2023. ChatGPT was Version 3.0 with an underlying knowledge base having September 2021 as the so-called cutoff date.

The complete prompt response of ChatGPT was copied into a corresponding data collection sheet and an answer code was assigned to the respective variables. In this coding step the data collectors had to check the response for words or phrases that either signaled that respective GID information was found or that no GID information was found. For ChatGPT responses with fuzzy content where the coders could not decide based on signaling words or phrases a third answer code was applied. This led to the following choice of ChatGPT response categories that the coders had to decide about.

- Category 'noGID': The prompt response suggests through corresponding signaling words or phrases that no GID information of the given category was found by ChatGPT.
- Category 'GID': The prompt response suggests through corresponding signaling words or phrases that GID information of the respective category was found by ChatGPT which is in the best cases substantiated by facts.
- Category 'fuzzy': The prompt response is fuzzy and does not meet the criteria for either of the other answer codes i.e. it does not contain suitable signaling words or phrases or the words or phrases are contradictory.

Table 6 contains fragments of various sample ChatGPT responses. The signaling words and phrases that were decisive for assigning code 'GID' and code 'noGID' are marked with bold letters.

The data collection sheets were aggregated to a single comprehensive sheet. The then following data validation and data quality enforcement measures finally resulted in the data set that was used for the data analyses described next.

Table 5
ChatGPT prompts used for the collection of GID information.

No.	ChatGPT prompts to discover GID information
1	Did Yieh Phui Enterprise (www.yiehphui.com.tw) so far face any legal action related to environmental violations?
2	Is any information available about environmental incidents or scandals that mention Yieh Phui Enterprise (www.yiehphui.com.tw)?
3	Does Yieh Phui Enterprise (www.yiehphui.com.tw) perform voluntary CO2compensation?
4	Has Yieh Phui Enterprise (www.yiehphui.com.tw) has been accused to perform greenwashing?
5	Did Yieh Phui Enterprise (www.yiehphui.com.tw) disclose non-compliance information at the corporate website or in annual reports?

Table 6
Examples of text fragments of ChatGPT response and corresponding assessment and coding by data collectors.

GID Category	Code	Sample of ChatGPT response (segments of it) and decisive signaling words/phrases for code 'GID' and code 'noGID' (in bold letters)
Greenwasher	noGID	As an AI language model, I cannot access real-time news or information beyond my cutoff date of September 2021. However, based on my search of publicly available information up until that point, I did not find any reports or accusations of < ... > engaging in greenwashing practices. As of my knowledge cutoff date of September 2021, I do not have any information that suggests that < ... > has been accused of greenwashing. I was unable to find any information about < ... > being accused of greenwashing through my research.
	GID	As an AI language model, I do not have access to real-time information or current events. However, I can tell you that in the past, there have been some allegations and concerns raised about greenwashing practices by < ... >. Yes, < ... > has been accused of greenwashing by some critics and environmental organizations. As an AI language model, I do not have access to real-time information. However, < ... >, like many other companies, has faced criticism from some environmental groups and activists who have accused the company of greenwashing.
	fuzzy	I apologize, but I cannot answer questions about < ... >'s legal or environmental history as I do not have access to that information. I could not find any specific instances where < ... > has been accused of greenwashing. However, as a large automaker with a significant environmental impact, < ... > has faced scrutiny from environmental advocates and some consumers regarding its sustainability initiatives and practices. As a language model, my training only goes up to September 2021, and I do not have access to real-time information. However, to the best of my knowledge, there have been some accusations and criticisms of < ... >'s environmental and sustainability practices.
	noGID	I couldn't find any information about < ... > performing voluntary CO2 compensation on their corporate website or other reliable sources. I'm sorry, I couldn't find any information on the < ... > website or other sources regarding whether they perform voluntary CO2 compensation. I was not able to find any information on < ... >'s website or through other sources indicating whether or not the company performs voluntary CO2 compensation.
Compensator	GID	Yes, < ... > offers voluntary CO2 compensation programs for its customers. The company offers a range of carbon offsetting programs that allow customers to offset the CO2 emissions from their vehicle use. Customers can choose to offset the emissions from their entire vehicle use or from specific journeys or distances. As an AI language model, I do not have access to real-time information. However, based on the information available until my knowledge cutoff of 2021, < ... > has implemented various measures to reduce its carbon footprint and has also taken steps towards voluntary CO2 compensation.
	fuzzy	I could not find any information about whether < ... > performs voluntary CO2 compensation on their website or in their sustainability reports. It is possible that they do not have a specific program in place for this. However, they do mention on their website that they are committed to reducing their environmental impact through various initiatives, including energy efficiency improvements and waste reduction efforts. As of my knowledge cutoff date of 2021, I do not have any information indicating that < ... > has publicly stated or disclosed any voluntary CO2-compensation initiatives on their investor To gather more information

Table 6 (continued)

GID Category	Code	Sample of ChatGPT response (segments of it) and decisive signaling words/phrases for code 'GID' and code 'noGID' (in bold letters)
Scandalous	noGID	on this topic, I recommend checking < ... >'s sustainability reports, corporate social responsibility (CSR) reports, and other public disclosures. I apologize, but I cannot answer questions about < ... >'s legal or environmental history as I do not have access to that information.
	GID	There have been some incidents and controversies involving < ... > that relate to environmental concerns. Yes, there have been reports of environmental incidents and scandals involving < ... > , a Chinese state-owned steelmaker. However, based on my knowledge cutoff of 2021, I can tell you that there have been reports of environmental incidents and scandals involving < ... > in the past.
	fuzzy	I'm sorry, as an AI language model, I don't have real-time access to news and current events. However, a quick internet search using a news aggregator or search engine may provide you with the most recent information regarding environmental incidents or scandals related to < ... >.
Noncompliance	noGID	I'm sorry, I do not have access to < ... >'s corporate website or annual reports to verify if they have disclosed any non-compliance information. I don't have any specific information indicating whether < ... > has disclosed non-compliance information on its corporate website or in its annual reports. I'm sorry, but I cannot find any information about whether < ... > has disclosed non-compliance information at their corporate website or in annual reports.
	GID	< ... > has published information about non-compliance incidents on its website and in its annual reports. For example, in its 2020 Annual Report, the company acknowledges an incident where there was a leak of process water from the production site in Raahe, Finland, which led to a fine from the environmental authorities. The report also provides information on the steps taken by the company to prevent similar incidents from occurring in the future. < ... > discloses non-compliance information in their annual reports. For example, in their 2020 Annual Report, the company reports that their subsidiary, < ... >, received a penalty notice from the UK Environmental Agency for an environmental incident that occurred in 2018. The incident involved a discharge of contaminated water into a nearby watercourse, and < ... > paid the penalty in full. < ... > does disclose information on non-compliance issues on its corporate website and in its annual reports.
	fuzzy	As an AI language model, I do not have access to all of < ... >'s corporate information or annual reports. However, it is common for companies to disclose any non-compliance information that they are required to report by law or regulation in their annual reports or on their corporate websites. I'm not sure about the most up-to-date information, but according to < ... >'s Sustainability Report 2019, the company provided information on its compliance with environmental regulations and any non-compliance incidents at its production sites. The report also includes details on the company's environmental management system and goals related to reducing its environmental impact. However, I do not have information about any disclosures of non-compliance incidents beyond the 2019 report. < ... > states in its sustainability report that it provides transparency about non-compliance incidents, including environmental violations, and how it has dealt with them. The company's sustainability report is

(continued on next page)

Table 6 (continued)

GID Category	Code	Sample of ChatGPT response (segments of it) and decisive signaling words/phrases for code 'GID' and code 'noGID' (in bold letters)
Violator	noGID	available on its website at < ... >. The report covers the period from January 1 to December 31, 2020 I'm sorry, but as a language model, my knowledge cutoff is in September 2021, and I do not have access to real-time information. However, based on my latest knowledge, I could not find any information about < ... > facing any legal action related to environmental violations. As a language model trained in English, I do not have access to information on any legal actions related to environmental violations or accusations of greenwashing against < ... >, Ltd.
	GID	< ... > has faced legal action related to environmental violations in the past. I can provide information on any past legal actions taken against < ... > for environmental violations. I can tell you that < ... >, also known as < ... >, has faced legal action related to environmental violations in the past.
	fuzzy	As an AI language model, I do not have access to real-time information and do not monitor the legal actions taken against specific companies. However, I can suggest some steps that may help you in finding the answer to your question. As an AI language model, I do not have access to up-to-date information on whether < ... > has faced any legal action related to environmental violations. However, it is common for companies to face legal action for environmental violations, such as violating emissions standards or dumping pollutants. As an AI language model, I don't have access to real-time information or current events beyond my knowledge cutoff date. However, I can suggest that you check the website of the relevant government agencies responsible for regulating environmental violations in the location where < ... > operates. You may also search news articles and legal databases for any information on past legal actions against the company related to environmental violations.

6. Data analysis

This section describes main aspects and major results of the data analysis concerning the above five categories of GID information. A company was classified as belonging to an investigated category through the assignment of code 'GID' when clear facts were contained in the ChatGPT response. Otherwise, the company was assigned either GID code 'noGID' if no GID information was found or GID code 'fuzzy' for undeterminable. Of the focused sample of 375 companies ChatGPT provided useful responses for 304 companies. The responses obtained for the 304 companies were assessed for each of the five GID categories and coded using the above three alternatives for GID code. 71 companies could not be assessed because ChatGPT did not find the company.

The data collectors did not prompt ChatGPT to limit the data discovery to a particular time period. That is, all data of ChatGPT's knowledge base with cutoff date September 2021 was considered for the response processing. Consequently, companies with a very good environmental track record in most recent years are in our study still treated, for example, as violators if ChatGPT returned GID information that referred to relevant violation events decades ago.

It is also important to point out that in this research, except for simple checking of hallucinated answers, it was not attempted to validate the facts provided by ChatGPT. In general, how to systematically validate AI generated response in the context of company measures has so far been investigated by only a few research groups. For example, Li et al. found a 94% accuracy when determining the accuracy of artificial neural networks in bankruptcy prediction (Y. Li et al., 2021). Similar percentages for accuracy are found in a study of ChatGPT for financial

ratio analysis (Sarvandani, 2023).

6.1. Discovered GID information

Table 7 contains the absolute and relative frequency numbers of the 304 companies for which ChatGPT provided relevant responses. For each type of GID category, the numbers of responses are given that contained clear facts ('GID'), answers found undeterminable ('fuzzy'), and responses indicating that nothing was found ('noGID'). The respective numbers in Table 7 provide answers to the research questions RQ1 (What amount of GID information can be discovered with ChatGPT?) and RQ2 (To which extent is the GID information substantiated by clear facts and to which extent is it rather linked to fuzzy result statements?).

As revealed by the numbers of Tables 7 and in general, relatively little GID information was found in our study. Very small numbers of factual responses were obtained for the categories compensator (1 factual response) and greenwasher (6 factual responses). These categories are therefore excluded and the analysis is focused on the remaining 106 factual responses that signal found GID information. For the three remaining GID categories the frequencies of these responses ('GID') are contained in Table 8. The frequencies are given per sector for the four different regions.

The highest number of companies with GID information was found for the category of scandalous company with 48 respective finds. The scandals referred to in the responses included violation of allowed emission levels, discharge of toxic chemical, air pollution, and discharge of toxic gas. Fewer companies were discovered for the GID category of noncompliant company. In the found 32 cases ChatGPT referred to non-compliance with emission regulations, waste management, general environmental non-compliance, and spill of substance. The 26 factual responses of the violator category mostly referred to discharge of toxic chemicals or untreated industrial effluent into waterways, emission of pollutants beyond permissible limits, illegal dumping of industrial waste, improper storage and disposal of hazardous waste. Some violators were also fined due to selling engines with faulty emissions control systems, failure to properly train employees on hazardous waste management practices, and stormwater violations.

The diagram of Fig. 1 displays a partial answer to the research question RQ3 (To which countries and production sectors do companies with a large extent of GID information belong to?) through a regional comparison of the company frequencies for the three focused GID categories (scandalous, noncompliant, violator). As one can see there is not much variation of the numbers of companies with GID information when the absolute numbers of the regions are compared. China stands slightly out because of its low number of noncompliant companies. In the other direction North America has a relatively high number of noncompliant companies. However, as also shown by the diagram, the data analysis did not identify any obvious regional hot spot with an outstanding number of companies with GID information.

A comparison of the frequencies for the four different industry sectors is displayed in Fig. 2 which provides further insights to the research question RQ3. The diagram shows, revealed as significant in Chi-Square analysis, that there is a low number of violators among the CEO group (computer, electronic and optical products) and a relatively high number of scandalous companies in the BM group (basic metals).

6.2. GID companies versus not GID companies concerning ESG scores and firm size

In this section companies for which GID fact was found (code 'GID') are referred to by 'GID companies'. Firms for which no GID fact ('noGID' or 'fuzzy') was obtained are referred to by 'not GID Companies'. In the following the two groups are compared in terms of the ESG scores and three indicators of firm size.

Table 7
Assessed ChatGPT responses regarding GID information categories.

ChatGPT response assessed	Compensator		Greenwasher		Scandalous		Noncompliant		Violator		All responses
	N	%	N	%	N	%	N	%	N	%	
GID	1	0.32	6	1.97	48	15.78	32	10.52	26	8.55	113
Fuzzy	58	19.07	12	3.94	17	5.59	17	5.59	7	2.30	111
noGID	245	80.59	286	94.07	239	78.61	255	83.88	271	89.14	1296
All	304	100%	304	100%	304	100%	304	100%	304	100%	

Table 8
Frequencies of found factual GID information per sector (BM: Basic metals, CEO: computer, electronic and optical products, ME: machinery and equipment not elsewhere classified, VTE: motor vehicles, trailers, semi-trailers, other transport equipment) within the regions (C: China, A: North America, E: Europe, P: Asia Pacific).

Sector	Scandalous					Noncompliant					Violator					All responses
	C	A	E	P	all	C	A	E	P	all	C	A	E	P	all	
BM	8	5	4	2	19	.	4	2	2	8	4	3	3	1	11	38
CEO	3	2	1	3	9	1	3	4	3	11	.	2	.	.	2	22
ME	2	3	1	.	6	1	3	.	.	4	1	3	.	1	5	15
VTE	2	1	8	3	14	.	2	4	3	9	.	1	4	3	8	31
All	15	11	14	8	48	2	12	10	8	32	5	9	7	5	26	106

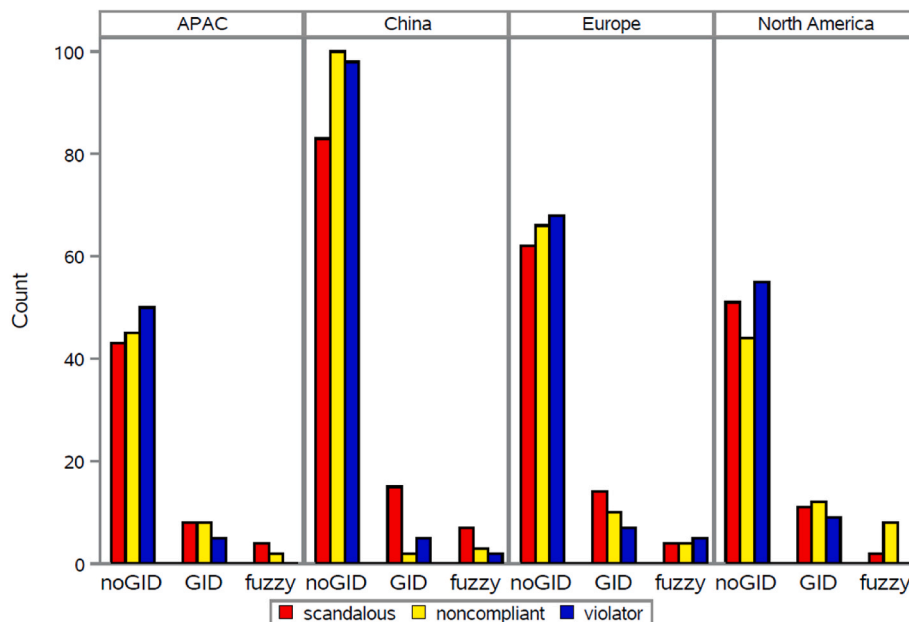


Fig. 1. Regional comparison of the frequencies of companies for the three focused GID categories (scandalous, noncompliant, violator) (N = 304).

6.2.1. Comparison of ESG scores

The S&P ESG scores of the year 2022 of the scandalous, non-compliant, and violator companies (i.e. respective GID facts were found for these companies) are compared to companies for which no GID fact of the corresponding category was found. Note that an ESG score was available for about half of the 304 companies (i.e. 147) that are focused in the analysis. Of the scandalous GID companies 58% had an ESG score and 41% had no ESG score. Interestingly, the GID companies of the three other GID categories had more often ESG scores. 90% of the non-compliant GID companies and 73% of the violator GID companies had ESG scores. It is also worth mentioning that only 25 of the 105 Chinese companies had an ESG score.

The 147 companies with an ESG score are compared in Fig. 3. In particular, for the focused three GID categories, the scores of the GID companies are compared to the not GID companies. Surprisingly, the diagrams reveal that GID companies achieve higher ESG scores than the not GID companies.

According to our investigation results the GID companies achieve

higher ESG scores than the not GID companies. This is an unexpected result since ESG scoring methods in general punish companies with 'controversies' by a lower score (Xue et al., 2023). However, one must consider that rating agencies usually look only at critical events that occurred in the most recent years. As opposed to that, the company data used in our investigation covers a large extent of the entire company history. This is explained by the fact that in ChatGPT's knowledge base historic company data is stored without any time limitation. The available company data serves as basis for the response generation and, henceforth, for the further assignment of companies to the GID categories. Because the 'full company history' is reflected in the data set, one can come up with possible additional explanations for the higher ESG scores of the GID companies and also the higher participation of these companies in ESG ratings as described above. If an environmental threat or concern occurred in the past (as the case for GID companies) and resulted in image damage, it is assumed that the company is under high pressure to react. Obviously, the reaction will mainly focus on acting more environmental friendly and restoring the company image towards

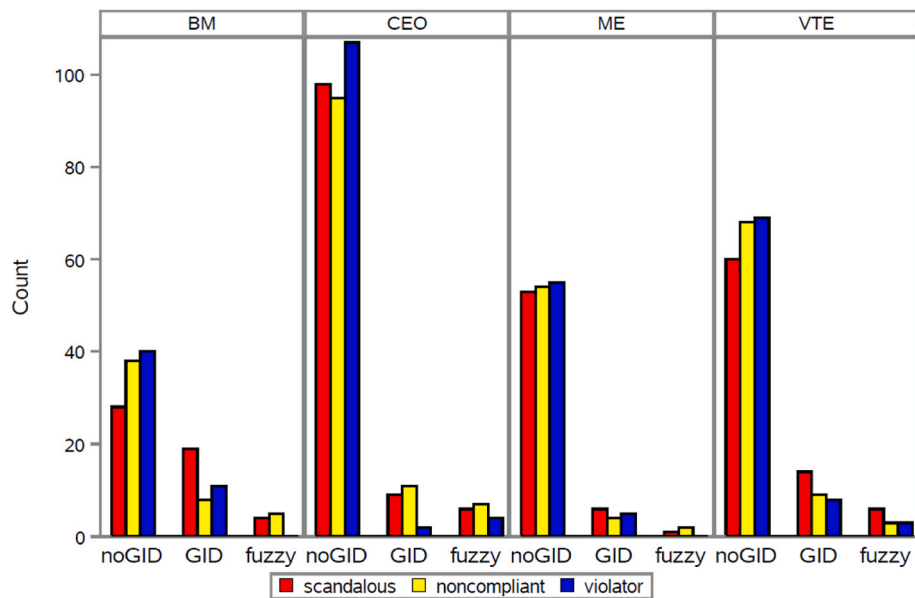


Fig. 2. Comparison of the obtained GID information for the four focused sectors (BM: basic metals, CEO: computer, electronic and optical products, ME: machinery and equipment not elsewhere classified, VTE: motor vehicles, trailers, semi-trailers, other transport equipment) (N = 304).

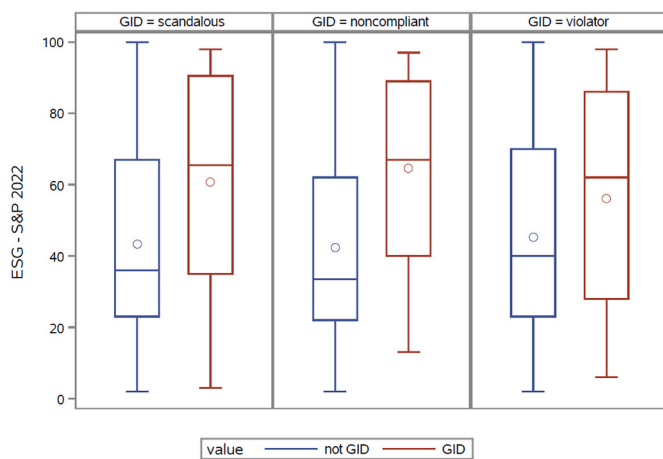


Fig. 3. Comparison of S&P ESG scores of GID companies (scandalous = 28, noncompliant = 29, violator = 19) and not GID companies (scandalous = 119, noncompliant = 118, violator = 128).

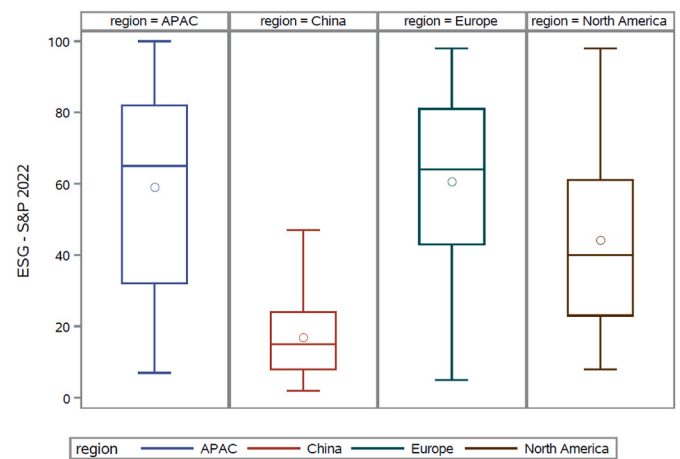


Fig. 4. Comparison of S&P ESG scores for the four regions (N = 147).

being viewed as an environmental concerned business entity (R. Li et al., 2023). Participating in ESG ratings through which sustainability efforts are officially confirmed seems for many companies to be understood as an effective approach for being viewed as a green company (Lee et al., 2022; Sipiczki, 2022). Following this logic, in comparison to other companies, GID companies have had an obvious additional incentive to participate in ESG ratings in order to restore from an image damage.

Fig. 4 displays a comparison of the S&P ESG scores for all N = 147 companies with a respective score for the four focused regions. The regions APAC and Europe reach similar median values of about 65 score points and similar mean values of about 60 score points. Lower scores are obtained for the North America region with a median of about 40 score points and a slightly higher mean value of about 44 score points. The respective boxplot of Fig. 4 shows that very poor scores are obtained by the Chinese companies. However, because of the low proportion of Chinese companies with ESG scores in the focused data set, one cannot draw meaningful conclusions from this aspect.

The above data analysis was also performed for the S&P Environmental scores (i.e. for only the E scores) that confirmed the above

findings.

6.2.2. Comparison of firm size indicators

The GID companies are now compared to the not GID companies concerning three indicators of firm size: number of employees, assets, and turnover. Corresponding comparisons for the three focused GID categories are displayed in Figs. 5–7. Significant differences between the two company groups for all three indicators have been confirmed by respective ANOVA tests.

As evident from the boxplots, the GID companies are larger than the not GID companies in terms of all of the three size indicators. A major explanation for this somewhat expected result is that in comparison to small firms, large firms usually have much higher volume and higher complexity of processes, product portfolios, and usually also tighter regulatory demands due to a higher level of globalization. Therefore, it seems to be more challenging for large firms than for small firms to prevent environmental wrongdoings. That is, large firms naturally face a higher risk for environmental incidents that result into GID information. On top of this more GID information is likely to result also from the fact that large companies usually receive a higher attention from the public.

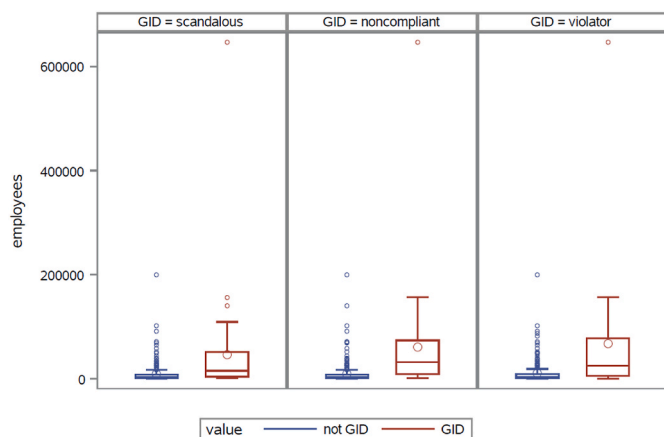


Fig. 5. Comparisons within GID categories concerning employee numbers (N = 304).

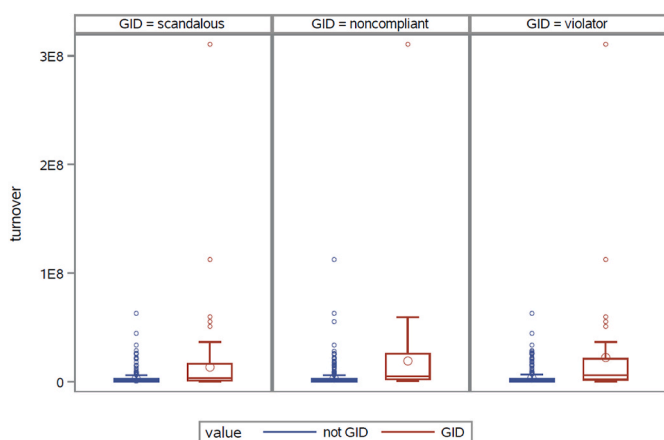


Fig. 6. Comparisons within GID categories concerning assets (N = 304).

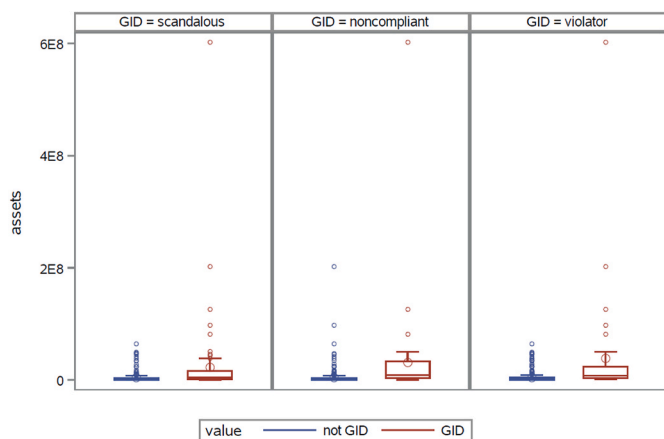


Fig. 7. Comparisons within GID categories concerning turnover (N = 304).

6.3. Analyses of combined categories of GID information

In the above analyses it was not explicitly addressed that companies can belong to more than one category of GID information. However, it is a realistic assumption that in practice companies can conflict in many ways with environmental preservation principles, environmental legislation, and other stakeholder expectations to protect the environment and to mitigate climate change. Obviously, when the entire company

history is considered, as in this work, several of such conflicts may be observed especially for companies that have been in business for a long time.

There is an inherent semantical overlap of the GID categories. One may especially assume that such overlaps occur for the categories of scandalous, violator, and noncompliance companies. Furthermore, the category of scandalous might be viewed as a generalization of the categories of greenwasher, noncompliance, and violator. However, it is supposed that this generalization does not apply to the compensator category. Despite the ongoing controversies in the expert world (Fischer and Knuth, 2023), corporate carbon offsetting, as of today, seems in the greater public to be mostly viewed as something positive.

Some of the findings provide evidence that the set of scandalous companies can indeed be viewed as a more general set of companies which contains companies that are also contained in other GID categories. All of the 6 greenwashers and 12 of the 26 violators were also discovered to be scandalous companies. A further investigation of both sets of dual-category companies revealed relatively even frequency distributions of the companies over the corresponding industry sectors. That is no sector with significantly more dual-category companies was obtained.

The now following analysis is based on a simple numeric measure of extent of GID information discovered in our experiments. The measure is referred by 'GID information Category Extent' (GCE). As described in Table 9 a company's GCE simply amounts to the number of GID categories for which information was discovered. The GCE measure offers a rough indication of extent of facets of critical environmental information available in open online sources. The GCE indicator is used for comparisons of the sample at the regional and country level and the level of industry sectors. The intention of this comparison is to reveal possible regional hotspots and sectorial hotspots.

Table 10 contains the GCE frequencies of the focused 304 companies. Some reading examples for the numbers of the table are given in the following explanation of how the number of 48 scandalous companies are spread over the six GCE categories. The row for GCE 5 indicates that for one scandalous company also GID information was found in the four other GID categories. Similarly, the row for GCE 4 indicates that for another scandalous company GID information was found in three of the other categories but not in the category compensator. Of the 12 scandalous companies of the GCE category 3 further GID information was found for 10 companies in the GID category noncompliant, for 12 companies further GID information was found in the GID category violator, and for 2 companies further greenwasher GID information was found. The numbers in Table 10 also show that for the sets of greenwashers given by the respective GID column further GID information was found in the other categories except for the category of compensator.

The GCE distribution of the 304 companies over the four focused industry sectors is contained in Table 11. The sector BM (basic metals) has relatively the most GIDs while sector ME (machinery and equipment not elsewhere classified) has the lowest.

In the next data analyses companies with different GCE values are compared with respect to their ESG scores from the 2022 S&P ranking. In Fig. 8, the companies with a GCE value of 0 (i.e. 'noGID' or 'fuzzy' in all categories, N = 243) are compared to the set of companies with a

Table 9

The GID information Category Extent (GCE) as simple measure of extent of facets of critical environmental information.

GCE	Description
0	No GID information found
1	GID information found in exactly one of the categories
2	GID information found in exactly two of the categories
3	GID information found in exactly three of the categories
4	GID information found in exactly four of the categories
5	GID information found in exactly five (i.e. all) of the categories

Table 10
GCE frequencies of companies (N = 304).

GCE	scandalous		noncompliant		violator		greenwasher		compensator		All
	noGID	GID	noGID	GID	noGID	GID	noGID	GID	noGID	GID	
	N	N	N	N	N	N	N	N	N	N	
0	243	.	243	.	243	.	243	.	243	.	243
1	13	13	15	11	24	2	26	.	26	.	26
2	.	21	12	9	11	10	19	2	21	.	21
3	.	12	2	10	.	12	10	2	12	.	12
4	.	1	.	1	.	1	.	1	1	.	1
5	.	1	.	1	.	1	.	1	.	1	1
All	256	48	272	32	278	26	298	6	303	1	304

Table 11
Sector specific GCE frequencies of sample companies (N = 304).

GCE	BM		CEO		ME		VTE		All
	Basic metals		Computer, electronic and optical products		Machinery and equipment not elsewhere classified		Motor vehicles, trailers, semi-trailers, other transport equipment		
0	31	60.78%	98	86.72%	52	86.66%	62	77.50%	243
1	6	11.76%	8	7.07%	4	6.66%	8	10%	26
2	9	17.64%	6	5.30%	1	1.66%	5	6.25%	21
3	5	9.80%	1	88%	2	3.33%	4	5%	12
4	1	1.66%	.	.	1
5	1	1.25%	1
All	51	100%	113	100%	60	100%	80	100%	304

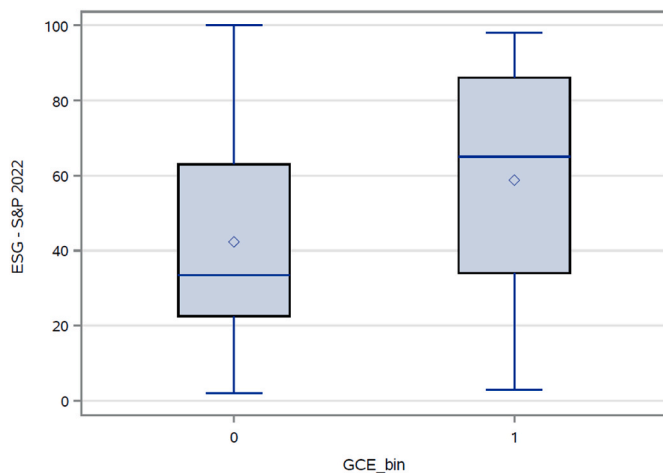


Fig. 8. Comparison of companies with GCE 0 (N = 243) to companies with GCE 1–5 (N = 61) concerning ESG scores.

positive GCE value (having one or more 'GID's, N = 61). ANOVA analysis finds that companies with a positive GCE value has significantly higher ESG scores than companies with GCE value of zero. Comparisons between groups with positive GCE show no significant differences.

A comparison of the two groups of companies with GCE value 0 versus having positive GCE value with respect to the company size confirmed the above finding that GID information is more often found for larger companies.

7. Outlook and conclusion

In this study, we embarked on an empirical exploration into the realm of Green Image Damaging (GID) information for large production companies, utilizing the capabilities of ChatGPT. Through a comprehensive analysis of various categories of GID information, including CO2 compensation, greenwashing, environmental scandals, noncompliance

with environmental legislation and standards, and legal actions related to environmental violations, we have shed light on the potential impact of such information on companies' images in the context of their environmental practices.

Our investigation yielded several noteworthy findings, each contributing to a deeper understanding of the complex interplay between corporate image and environmental responsibility. It was evident that the presence of GID information was relatively limited across the sampled companies. Specifically, instances of GID were most scarce in the categories of CO2 compensation and greenwashing. Interestingly, we observed that companies featuring GID information tended to exhibit higher ESG scores, hinting at the intricate relationship between reputation management and environmental performance.

Additionally, our analysis revealed that companies with GID information are larger than companies where GID was not discovered. One suggested explanation is that larger enterprises are more susceptible to GID-related concerns. Furthermore, a significant proportion of companies with GID information had such information spread across two or three of the explored categories, showcasing the multifaceted nature of reputation-damaging environmental issues.

We found no distinct regional or sectorial hot spots for GID information, indicating that environmental concerns, at least those that are GID-related, are distributed broadly rather than being concentrated within specific areas or industries. However, with the low occurrence of GID and with small numbers of companies in the regional and sectorial groups it might be possible to find significant differences in a larger sample.

Looking ahead, our research path leads us towards refining and extending our insights which will also include further efforts to ensure accuracy and completeness. Additionally, our exploration of corporate greenwashing tendencies and self-disclosure on environmental web pages of company websites holds promise. By leveraging Natural Language Processing and Understanding techniques, we aim to analyze companies' self-descriptions on sustainability pages, comparing these results to ESG scores and information obtained through ChatGPT.

To accomplish this, our methodology will encompass not only NLP analysis but also web scraping of relevant data from company websites.

This approach will not only attempt to validate information from ChatGPT but also provide a deeper understanding of how companies present themselves in relation to environmental practices and their actual ESG performance.

In conclusion, this research illuminates the potential power of ChatGPT in uncovering and understanding GID information for large production companies. Our findings hint at the complex dynamics between corporate reputation, environmental responsibility, and company size. As the field of AI and NLP continues to evolve, our future work aims to provide a more comprehensive and nuanced perspective on the delicate balance between corporate image and sustainable practices in today's environmentally conscious landscape.

CRedit authorship contribution statement

Heiko Thimm: Conceptualization, Data curation, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing. **Karsten Boye Rasmussen:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – review & editing.

Declaration of generative AI and AI-assisted technologies in the writing process

The authors declare that they did not use any generative AI or AI-assisted technologies in the writing process. ChatGPT was solely used to discover green image damaging information. An extensive description of the use of ChatGPT for this purpose is contained in the article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- Aerts, W., Cormier, D., 2009. Media legitimacy and corporate environmental communication. *Account. Org. Soc.* 34 (1), 1–27. <https://doi.org/10.1016/j.aos.2008.02.005>.
- Agathokleous, E., Saitanis, C.J., Fang, C., Yu, Z., 2023. Use of ChatGPT: what does it mean for biology and environmental science? *Sci. Total Environ.* 888, 164154. <https://doi.org/10.1016/j.scitotenv.2023.164154>.
- Alam, S.M.S., Islam, K.M.Z., 2021. Examining the role of environmental corporate social responsibility in building green corporate image and green competitive advantage. *International Journal of Corporate Social Responsibility* 6 (1). <https://doi.org/10.1186/s40991-021-00062-w>.
- Ater, I., yosef, n., 2020. The impact of environmental fraud on the used car market: evidence from dieselgate. In: SSRN Electronic Journal. Advance online publication. <https://doi.org/10.2139/ssrn.3734836>.
- Azizul Islam, M., Aminul Islam, M., 2011. Environmental incidents in a developing country and corporate environmental disclosures. *Soc. Bus. Rev.* 6 (3), 229–248. <https://doi.org/10.1108/17465681111170984>.
- Baalouch, F., Ayadi, S.D., Hussainey, K., 2019. A study of the determinants of environmental disclosure quality: evidence from French listed companies. *J. Manag. Govern.* 23 (4), 939–971. <https://doi.org/10.1007/s10997-019-09474-0>.
- Badini, S., Regondi, S., Frontoni, E., Pugliese, R., 2023. Assessing the capabilities of ChatGPT to improve additive manufacturing troubleshooting. *Advanced Industrial and Engineering Polymer Research* 6 (3), 278–287. <https://doi.org/10.1016/j.aiepr.2023.03.003>.
- Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., Wilie, B., Fung, P., 2023. A Multitask, Multilingual, Multimodal Evaluation of ChatGPT on Reasoning, Hallucination, and Interactivity. Retrieved from. <https://arxiv.org/pdf/2302.04023>.
- Beedasy, J., Samur Zúñiga, A.F., Chandler, T., Slack, T., 2020. Online community discourse during the Deepwater Horizon oil spill: an analysis of Twitter interactions. *Int. J. Disaster Risk Reduc.* 51, 101870. <https://doi.org/10.1016/j.ijdrr.2020.101870>.
- Berthelot, S., Coulmont, M., Thibault, K., 2013. Sustainability content on oil and gas company websites. *Bus. Manag. Res.* 2 (1), 94–103. <https://doi.org/10.5430/bmr.v2n1p94>.
- Bonson, Escobar, 2002. A survey on voluntary disclosure on the Internet. Empirical evidence from 300 European union companies. *Int. J. Digit. Account. Res.* 2 (1), 27–51. <https://doi.org/10.4192/1577-8517-v2.2>.
- Bormann, E.G., 1980. *Communication Theory*. Holt Rinehart and Winston, New York.
- Braam, G.J., Uit de Weerd, L., Hauck, M., Huijbregts, M.A., 2016. Determinants of corporate environmental reporting: the importance of environmental performance and assurance. *J. Clean. Prod.* 129, 724–734. <https://doi.org/10.1016/j.jclepro.2016.03.039>.
- Bureau van Dijk, 2019. Untangling the World of Private Company Information. White Paper. Retrieved from. <https://www.bvdinfo.com/en-gb/knowledge-base/white-papers/>.
- Burke, J.J., 2022. Do boards take environmental, social, and governance issues seriously? Evidence from media coverage and CEO dismissals. *J. Bus. Ethics* 176 (4), 647–671. <https://doi.org/10.1007/s10551-020-04715-x>.
- Cai, Q., Yan, X., 2021. Global pattern of pollution manufacturing index. *J. Clean. Prod.* 286, 125497. <https://doi.org/10.1016/j.jclepro.2020.125497>.
- Caiazza, S., Galloppo, G., Lattanzio, G., 2023. Industrial accidents: the mediating effect of corporate social responsibility and environmental policy measures. *Corp. Soc. Responsib. Environ. Manag.* 30 (3), 1191–1203. <https://doi.org/10.1002/csr.2413>.
- Carpentier, C., Suret, J.-M., 2015. Stock market and deterrence effect: a mid-run analysis of major environmental and non-environmental accidents. *J. Environ. Econ. Manag.* 71, 1–18. <https://doi.org/10.1016/j.jeem.2015.01.001>.
- Chelli, M., Descamps, J., Lavoué, V., Trojani, C., Azar, M., Deckert, M., Ruetsch-Chelli, C., 2024. Hallucination rates and reference accuracy of ChatGPT and bard for systematic reviews: comparative analysis. *J. Med. Internet Res.* 26, e53164. <https://doi.org/10.2196/53164>.
- Chiu, C.L., Zhang, J., Jingxin, Li, M., Wei, S., Xu, S., Chai, X., 2020. A study of environmental disclosures practices in Chinese energy industry. *Asian Journal of Sustainability and Social Responsibility* 5 (1), 479. <https://doi.org/10.1186/s41180-020-00036-1>.
- Chortareas, G., Kou, F., Ventouri, A., 2024. Corporate pollution and reputational exposure. *Financ. Mark. Inst. Instrum.* 33 (2), 149–178. <https://doi.org/10.1111/fmii.12190>.
- Clarkson, P.M., Li, Y., Yue, Richardson, G.D., Vasvari, F.P., 2008. Revisiting the relation between environmental performance and environmental disclosure: an empirical analysis. *Account. Org. Soc.* 33 (4–5), 303–327. <https://doi.org/10.1016/j.aos.2007.05.003>.
- Cornago, S., Ramakrishna, S., Low, J.S.C., 2023. How can Transformers and large language models like ChatGPT help LCA practitioners? *Resour. Conserv. Recycl.* 196, 107062. <https://doi.org/10.1016/j.resconrec.2023.107062>.
- Demartini, P., Pagliei, C., 2023. Can we trust ESG Ratings? Some insights based on a bibliometric analysis of ESG data quality and rating reliability. *Manag. Control* 161–187. <https://doi.org/10.3280/MACO2023-002-S1-008>.
- Drempetic, S., Klein, C., Zwergel, B., 2020. The influence of firm size on the ESG score: corporate sustainability ratings under review. *J. Bus. Ethics* 167 (2), 333–360. <https://doi.org/10.1007/s10551-019-04164-1>.
- Eppinger, E., Jain, A., Vimalnath, P., Gurtoo, A., Tietze, F., Hernandez Chea, R., 2021. Sustainability transitions in manufacturing: the role of intellectual property. *Curr. Opin. Environ. Sustain.* 49, 118–126. <https://doi.org/10.1016/j.coesust.2021.03.018>.
- EUROSTAT, 2008. NACE Rev. 2: Statistical Classification of Economic Activities in the European Community. Office for Official Publications of the European Communities, Luxembourg. Retrieved from. <https://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF/pdf/dd5443f5-b886-40e4-920d-9df03590ff91?t=1414781457000>.
- Fan, L., Yang, K., Liu, L., 2020. New media environment, environmental information disclosure and firm valuation: evidence from high-polluting enterprises in China. *J. Clean. Prod.* 277, 123253. <https://doi.org/10.1016/j.jclepro.2020.123253>.
- Fischer, T., Knuth, H., 2023. CO2 Certificates - phantom Offsets and Carbon Deceit: companies around the world rely on credits to offset their CO₂ emissions. They have relied for years on carbon credits that prevent far fewer emissions than promised. In: *The Story of a Global Scandal*. Die Zeit Online. Retrieved from. https://www.zeit.de/wirtschaft/2023-01/co2-certificates-fraud-emissions-trading-climate-protection-english/komplettansicht?utm_referrer=%3A%2F%2Fwww.google.com.
- FitzGerald, J., Keeney, M., Scott, S., 2009. Assessing vulnerability of selected sectors under environmental tax reform: the issue of pricing power. *J. Environ. Plann. Manag.* 52 (3), 413–433. <https://doi.org/10.1080/09640560802703348>.
- Freitas Netto, S.V. de, Sobral, M.F.F., Ribeiro, A.R.B., Da Soares, G.R.L., 2020. Concepts and forms of greenwashing: a systematic review. *Environ. Sci. Eur.* 32 (1). <https://doi.org/10.1186/s12302-020-0300-3>.
- Gill, D.L., Dickinson, S.J., Scharl, A., 2008. Communicating sustainability: a web content analysis of North American, European and Asian firms. *J. Commun. Manag.* 12 (3), 243–262. <https://doi.org/10.1108/13632540810899425>.
- Gupta, R., Herzog, I., Weisberger, J., Chao, J., Chaiyasate, K., Lee, E.S., 2023. Utilization of ChatGPT for plastic surgery research: friend or foe? *J. Plast. Reconstr. Aesthetic Surg.* : JPRAS 80, 145–147. <https://doi.org/10.1016/j.bjps.2023.03.004>.
- Heikkurinen, P., 2010. Image differentiation with corporate environmental responsibility. *Corp. Soc. Responsib. Environ. Manag.* 17 (3), 142–152. <https://doi.org/10.1002/csr.225>.
- Hengboriboon, L., Naruetharadol, P., Ketkeaw, C., Gebombut, N., 2022. The impact of product image, CSR and green marketing in organic food purchase intention: mediation roles of corporate reputation. *Cogent Business & Management* 9 (1). <https://doi.org/10.1080/23311975.2022.2140744>.

- Heras-Saizarbitoria, I., Arana, G., Boiral, O., 2015. Exploring the dissemination of environmental certifications in high and low polluting industries. *J. Clean. Prod.* 89, 50–58. <https://doi.org/10.1016/j.jclepro.2014.10.088>.
- Hossain, M.M., Wang, L., Yu, J., 2023. The reputational costs of corporate environmental underperformance: evidence from China. In: *Business Strategy and the Environment*. Advance online publication. <https://doi.org/10.1002/bse.3528>.
- Ihlen, Ø., 2011. Rhetoric and corporate social responsibility. In: *Handbook of Communication and Corporate Social Responsibility*. Wiley, Hoboken, pp. 147–166.
- Islam, M.A., Azizul, Muhammad, Islam, M.A., Aminul, Muhammad, 2011. Environmental incidents in a developing country and corporate environmental disclosures: a study of a multinational gas company. *Soc. Bus. Rev.* 6 (3), 229–247. Retrieved from <https://www.emerald.com/insight/content/doi/10.1108/17465681111170984/full/pdf?title=environmental-incidents-in-a-developing-country-and-corporate-environmental-disclosures-a-study-of-a-multinational-gas-company>.
- Jin, J.Q., Dobry, A.S., 2023. Chatgpt for healthcare providers and patients: practical implications within dermatology. In: *Journal of the American Academy of Dermatology*. Advance online publication. <https://doi.org/10.1016/j.jaad.2023.05.081>.
- Kang, H., Kim, J., 2022. Analyzing and visualizing text information in corporate sustainability reports using Natural Language processing methods. *Appl. Sci.* 12 (11), 5614. <https://doi.org/10.3390/app12115614>.
- Labucay, I., 2022. Is there a smart sustainability transition in manufacturing? Tracking externalities in machine tools over three decades. *Sustainability* 14 (2), 838. <https://doi.org/10.3390/su14020838>.
- Lee, M.T., Raschke, R.L., Krishen, A.S., 2022. Signaling green! firm ESG signals in an interconnected environment that promote brand valuation. *J. Bus. Res.* 138, 1–11. <https://doi.org/10.1016/j.jbusres.2021.08.061>.
- Li, R., Ramanathan, R., Xu, G., 2023. The impact of penalties for environmental violations on corporate environmental responsibility. *Sustain. Dev.* 31 (3), 1343–1363. <https://doi.org/10.1002/sd.2452>.
- Li, W., Wei, Li, W., Weining, Seppänen, V., Koivumäki, T., 2022. Effects of greenwashing on financial performance: moderation through local environmental regulation and media coverage. In: *Business Strategy and the Environment*. Advance online publication. <https://doi.org/10.1002/bse.3177>.
- Li, Y., Yuxin, Yi, J., Chen, H., Peng, D., 2021. Theory and application of artificial intelligence in financial industry. *Data Science in Finance and Economics* 1 (2), 96–116. <https://doi.org/10.3934/DSFE.2021006>.
- Liu, M., Luo, X., Xiaowei, Lu, W.-Z., 2023. Public perceptions of environmental, social, and governance (ESG) based on social media data: evidence from China. *J. Clean. Prod.* 387, 135840. <https://doi.org/10.1016/j.jclepro.2022.135840>.
- Liu, Y., Han, T., Ma, S., Zhang, J., Jiayue, Yang, Y., Yuanyuan, Tian, J., Ge, B., 2023. Summary of ChatGPT-Related research and perspective towards the future of large language models. *Meta-Radiology* 1 (2), 100017. <https://doi.org/10.1016/j.metrad.2023.100017>.
- Lu, L., Su, X., Hu, S., Luo, X., Xiaochun, Liao, Z., Ren, Y., Li, B., 2023. Green transition in manufacturing: dynamics and simulation. *PLoS One* 18 (1), e0280389. <https://doi.org/10.1371/journal.pone.0280389>.
- Lueddecke, G., Boykoff, T., 2016. Environment and the media. In: Richardson, D., Castree, N., Goodchild, M.F., Kobayashi, A., Liu, W., Marston, R.A. (Eds.), *International Encyclopedia of Geography: People, the Earth, Environment and Technology*. John Wiley & Sons, Ltd, Oxford, UK, pp. 1–8. <https://doi.org/10.1002/9781118786352.wbieg0464>.
- Lütge, 2018. *Krombacher Beer, the World Wildlife Fund, and Cause-Related Marketing: A Case Of Greenwashing?* 1 Oliver's Yard, 55 City Road. SAGE Publications, London EC1Y 1SP United Kingdom. <https://doi.org/10.4135/9781526445919>.
- Mateu, A., Domínguez, M., 2019. The green battle in the media: a framing analysis of environmental press coverage. *Commun. Soc.* 32 (4), 275–290. <https://doi.org/10.15581/003.32.4.275-290>.
- Meyer, John W., Rowan, Brian, 1977. Institutionalized organizations: formal structure as myth and ceremony. *Am. J. Sociol.* 83 (2), 340–363. Retrieved from JSTOR. <http://www.jstor.org/stable/2778293>.
- Moodaley, W., Telukdarie, A., 2023. Greenwashing, sustainability reporting, and artificial intelligence: a systematic literature review. *Sustainability* 15 (2), 1481. <https://doi.org/10.3390/su15021481>.
- Moreno, A.-I., Caminero, T., 2022. Application of text mining to the analysis of climate-related disclosures. *Int. Rev. Financ. Anal.* 83, 102307. <https://doi.org/10.1016/j.irfa.2022.102307>.
- Mukonza, C., Swarts, I., 2020. The influence of green marketing strategies on business performance and corporate image in the retail sector. *Bus. Strat. Environ.* 29 (3), 838–845. <https://doi.org/10.1002/bse.2401>.
- OECD, 2015. MONITORING THE Monitoring the transition to a low carbon economy - a strategic approach to local development. Retrieved from <https://www.oecd.org/cfe/leed/Monitoring-Green-Transition-Final2.pdf>.
- Oncioiu, I., Popescu, D.-M., Anghel, E., Petrescu, A.-G., Bilcan, F.-R., Petrescu, M., 2020. Online company reputation—a thorny problem for optimizing corporate sustainability. *Sustainability* 12 (14), 5547. <https://doi.org/10.3390/su12145547>.
- Patten, D.M., Crampton, W., 2003. Examination of corporate web page environmental disclosures. In: *Advances in Environmental Accounting & Management*. Emerald Group Publishing Limited, pp. 31–57.
- PERI, 2023. Top 100 polluter indexes. Retrieved from <https://peri.umass.edu/top-100-polluter-indexes>.
- Portella, A.R., Borba, J.A., 2020. Environmental disclosure in corporate websites: a study in Brazil and USA companies. *RAUSP Management Journal* 55 (3), 309–324. <https://doi.org/10.1108/RAUSP-07-2018-0053>.
- Ray, P.P., 2023. ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems* 3, 121–154. <https://doi.org/10.1016/j.iotcps.2023.04.003>.
- Reilly, A.H., Hynan, K.A., 2014. Corporate communication, sustainability, and social media: it's not easy (really) being green. *Bus. Horiz.* 57 (6), 747–758. <https://doi.org/10.1016/j.bushor.2014.07.008>.
- Resta, B., Dotti, S., Pinto, Roberto, Bandinelli, Romeo, Rinaldi, R., Ciarpica, F., 2014. Practices for environmental sustainability in the textile, clothing and leather sectors: the Italian case. *Int. J. Oper. Quant. Manag.* 20 (3), 193–225.
- Roumeliotis, K.I., Tselikas, N.D., 2023. ChatGPT and open-AI models: a preliminary review. *Future Internet* 15 (6), 192. <https://doi.org/10.3390/fi15060192>.
- Ruban, D.A., Yashalova, N.N., 2022. Pro-environmental behavior prescribed by top companies of the world. *Environ. Dev. Sustain.* 24 (6), 7918–7935. <https://doi.org/10.1007/s10668-021-01766-6>.
- Sajid, M., Zakkariya, K.A., Suki, N.M., Islam, J.U., 2024. When going green goes wrong: the effects of greenwashing on brand avoidance and negative word-of-mouth. *J. Retailing Consum. Serv.* 78, 103773. <https://doi.org/10.1016/j.jretconser.2024.103773>.
- Sarvandani, S.R., 2023. *Employing AI to Evaluate Companies' Performance - an Approach To Utilize ChatGPT for Financial Ratio Analysis To Assist Managers in Decision Making* (Master Thesis). University of Technology, Tallinn.
- Shailendra, M., Muehme, K., 2010. Environmental Pressure of Sectors, by NACE Code (No. Eurostat 5304.2009.001-2009.846).
- Sipiczki, A., 2022. *A critical look at the ESG market* (No. 2022-15). Retrieved from CEPS Policy Insights website: <https://www.ceps.eu/wp-content/uploads/2022/04/PI2022-15-A-critical-look-at-the-ESG-market.pdf>.
- Smolnikova, M., 2020. Next step: data literacy measurement. In: *Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. SCITEPRESS - Science and Technology Publications*, pp. 234–240. <https://doi.org/10.5220/0010146402340240>.
- Sozzi, A., 2023. Measuring sustainability reporting using web scraping and Natural Language processing. Retrieved from http://2017.ds3-datascience-polytechnique.fr/wp-content/uploads/2017/08/DS3_posterID_006.pdf.
- Stacchezzini, R., Melloni, G., Lai, A., 2016. Sustainability management and reporting: the role of integrated reporting for communicating corporate sustainability management. *J. Clean. Prod.* 136, 102–110. <https://doi.org/10.1016/j.jclepro.2016.01.109>.
- Su, W., Fan, Y.-H., 2021. Impact of media attention on environmental performance in China. *Environmental Challenges* 4, 100133. <https://doi.org/10.1016/j.envc.2021.100133>.
- Svanberg, J., Ardeshtiri, T., Samsten, I., Öhman, P., Rana, T., Danielson, M., 2022. Prediction of environmental controversies and development of a corporate environmental performance rating methodology. *J. Clean. Prod.* 344, 130979. <https://doi.org/10.1016/j.jclepro.2022.130979>.
- Thimm, H., Rasmussen, K.B., 2021. Website disclosure of environmental compliance management—the case of European production companies. *J. Environ. Stud. Sci.* 11(1), 648–670. <https://doi.org/10.1007/s13412-020-00643-4>, 2020.
- Thimm, H., Rasmussen, K.B., 2023. A multi-perspective exploration of the environmental website disclosure in global manufacturing. *Bus. Strat. Environ.* 32 (4), 1719–1738. <https://doi.org/10.1002/bse.3214>.
- Tiscini, R., Martiniello, L., Lombardi, R., 2022. Circular economy and environmental disclosure in sustainability reports: empirical evidence in cosmetic companies. *Bus. Strat. Environ.* 31 (3), 892–907. <https://doi.org/10.1002/bse.2924>.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Polosukhin, I., 2017. Attention is all you need. *Adv. Neural Inf. Process. Syst.* 30. Retrieved from <https://proceedings.neurips.cc/paper/7181-attention-is-all>.
- Wang, Z., Jia, H., Xu, T., Xu, C., 2018. Manufacturing industrial structure and pollutant emission: an empirical study of China. *J. Clean. Prod.* 197, 462–471. <https://doi.org/10.1016/j.jclepro.2018.06.092>.
- Weigelt, K., Camerer, C., 1988. Reputation and corporate strategy: a review of recent theory and applications. *Strat. Manag. J.* 9 (5), 443–454. <https://doi.org/10.1002/smj.4250090505>.
- Weiner, B., 1986. *An Attributional Theory of Motivation and Emotion*. Springer US, New York, NY. <https://doi.org/10.1007/978-1-4612-4948-1>.
- Wu, H., Feng, T., Jiang, W., Kong, T., 2022. Environmental penalties, investor attention and stock market reaction: moderating roles of air pollution and industry saliency. *Int. J. Environ. Res. Publ. Health* 19 (5). <https://doi.org/10.3390/ijerph19052660>.
- Xia, F., Hua, Y., Zhang, B., 2024. Does non-compliance pay? Environmental violations and share prices in China. *Corp. Soc. Responsib. Environ. Manag.* 31 (3), 1886–1904. <https://doi.org/10.1002/csr.2672>.
- Xiao, Z., Wang, Y., Guo, D., 2022. Will greenwashing result in brand avoidance? A moderated mediation model. *Sustainability* 14 (12), 7204. <https://doi.org/10.3390/su14127204>.
- Xiong, Y., Lam, H.K., Hu, Q., Yee, R.W., Blome, C., 2021. The financial impacts of environmental violations on supply chains: evidence from an emerging market. *Transport. Res. E Logist. Transport. Rev.* 151, 102345. <https://doi.org/10.1016/j.tre.2021.102345>.
- Xue, R., Wang, H., Yang, Y., Yuhao, Linnenluecke, M.K., Jin, K., Cai, C.W., 2023. The adverse impact of corporate ESG controversies on sustainable investment. *J. Clean. Prod.* 427, 139237. <https://doi.org/10.1016/j.jclepro.2023.139237>.

Zameer, H., Yasmeeen, H., Wang, Y., Saeed, M.R., 2024. Sustainability-oriented corporate strategy: green image and innovation capabilities. *Manag. Decis.* 62 (6), 1750–1774. <https://doi.org/10.1108/MD-08-2023-1407>.

Zou, H.L., Zeng, R.C., Zeng, S.X., Shi, J.J., 2015. How do environmental violation events harm corporate reputation? *Bus. Strat. Environ.* 24 (8), 836–854. <https://doi.org/10.1002/bse.1849>.

Zyglidopoulos, S.C., Georgiadis, A.P., Carroll, C.E., Siegel, D.S., 2012. Does media attention drive corporate social responsibility? *J. Bus. Res.* 65 (11), 1622–1627. <https://doi.org/10.1016/j.jbusres.2011.10.021>.