Enhancing Sustainability Reporting through Automated Text Analysis: A Systematic Review and Empirical Study in the Apparel Industry

Mahsa Mohammadrezaei

Thesis submitted to the University of Ottawa

in partial Fulfillment of the requirements for the

PhD in Digital Transformation and Innovation

Faculty of Engineering

University of Ottawa

Acknowledgments

As I reach the end of this journey, I am overwhelmed with gratitude for the incredible support and encouragement I have received along the way. This thesis is not just a reflection of my work but a testament to the love, guidance, and inspiration of the many people who have stood by me.

First and foremost, I would like to express my deepest appreciation to my advisor, Professor Jose Carlos Marques. Your wisdom, patience, and unwavering belief in my abilities have been my guiding light. You have not only shaped my academic growth but also taught me the true essence of perseverance and passion. I will forever be grateful for the countless hours you spent mentoring me, the thoughtful feedback you provided, and the moments of reassurance you offered when I needed them most.

I am deeply indebted to my family, whose love and support have been my anchor. To the five most important people in my life—my parents, Mozhgan and Hossein; my husband, Arang; my beloved sister, Dorsa; and Hossein—your sacrifices and unwavering belief in me have been the foundation upon which I built my dreams. Your endless encouragement has given me the courage to pursue this path, and I am forever grateful for the values you instilled in me.

This thesis is a reflection of not just my work, but the collective support, love, and wisdom of everyone who has been a part of this journey. I am profoundly grateful to each and every one of you.

Thank you, from the bottom of my heart.

Table of Contents

Enhancing Sustainability Reporting through Automated Text Analysis: A Systematic Review and Empirical Study in the Apparel Industry

Abstract- This thesis investigates the transformative potential of automated text analysis techniques, specifically text mining (TM) and natural language processing (NLP), in advancing the analysis of sustainability reporting (SR). Through two interrelated studies, critical gaps in the field are addressed by evaluating the current application of text analysis to SR analysis and introducing a novel analytical framework. The first study offers a comprehensive systematic literature review (SLR) of TM and NLP applications in SR, assessing methodologies, research objectives, and analytical depth. The findings underscore both the promise and limitations of automated text analysis in extracting meaningful insights from SRs, highlighting untapped potential. Building on these insights, the second study applies advanced text mining techniques to a case study within the apparel industry, focusing on the years following the Rana Plaza disaster. This application reveals biases toward positive reporting, raising concerns about the transparency and credibility of SRs. Together, these studies enhance the theoretical understanding and practical application of TM-NLP tools in SR, advocating for more transparent, balanced, and credible sustainability practices.

Keywords: systematic literature review, sustainability reports, text mining, natural language processing, NLP, big data, SASB, GRI, BERT, Sentiment analysis

PhD Thesis

University of Ottawa, Ottawa, Canada

Fall 2024

1 Introduction

This thesis is composed of two interrelated studies, each structured as a scholarly article, that together delve into the transformative role of automated text analysis in sustainability reporting (SR). The research investigates the implications and potential of these advanced technologies in enhancing the analysis and communication of sustainability data. The first study provides a comprehensive systematic literature review of text mining (TM) and natural language processing (NLP) techniques as they are applied to the analysis of SRs. The second study builds on this foundation by examining the practical implementation of these automated methods, focusing specifically on their application in conducting materiality assessments and balance analysis within SRs, particularly within the context of the apparel industry and the aftermath of the Rana Plaza disaster.

This introductory chapter is structured as follows: Section 1.1 outlines the challenges in sustainability reporting, particularly focusing on the complexities and limitations inherent in current practices. It emphasizes the critical need for more transparent, comprehensive, and accurate reporting, highlighting the growing concerns around the quality and credibility of sustainability reports (SRs). Section 1.2 identifies significant research gaps, focusing on the underexplored intersection of materiality and balance in SRs. Section 1.3 presents the two key papers that form the core of this thesis and the overall thesis contribution.

1.1 Challenges in Sustainability Reporting

In recent years, the rise in corporate disclosure of non-financial data related to social and environmental issues—commonly referred to as sustainability reporting (SR)—has been accompanied by growing concerns regarding the quality, transparency, and credibility of these reports (Lindgren et al., 2021). Although the volume has increased, there are concerns about the quality and transparency of these disclosures (Aras & Crowther, 2008; Billio et al., 2021; Boiral & Heras-Saizarbitoria, 2020; Bushee et al., 2018; Clarkson et al., 2019; Fabrizio & Kim, 2019). A significant challenge within SR is the complex interplay between materiality and balance, which are critical for ensuring that the information disclosed is both relevant and trustworthy. Additionally, the sheer volume, diversity, and complexity of the data within SRs pose significant

challenges for evaluation, making analysis, interpretation, and comparison difficult (Michelon et al., 2015, 2022).

Materiality and balance are cornerstone principles in sustainability reporting, each playing a critical role in ensuring that corporate disclosures are relevant, transparent, and credible. Materiality focuses on identifying the environmental, social, and governance (ESG) issues that are most significant to both the company and its stakeholders. These issues are those that could materially impact a company's performance, reputation, and ability to create long-term value (Dhaliwal et al., 2011; Lai & Stacchezzini, 2021). Companies are expected to engage with stakeholders to prioritize these material issues in their reports, reflecting the most critical aspects that influence decision-making (Baumüller & Sopp, 2021; Cerbone & Maroun, 2020; Henriques et al., 2022). However, determining what is considered material remains a subjective process that differs significantly across companies, resulting in inconsistencies and selective reporting. Leading frameworks like the Sustainability Accounting Standards Board (SASB) and the Global Reporting Initiative (GRI) tackle these issues by offering guidelines and standards to help define materiality in ESG reporting. However, tackling materiality is just one aspect of the challenge. Balance, on the other hand, ensures that sustainability reports provide an objective and comprehensive view of a company's sustainability performance. Balanced reporting requires transparency in presenting both positive achievements and challenges, thus avoiding the tendency to use sustainability reports merely as promotional tools. This balanced approach is essential for building trust with stakeholders and demonstrating a genuine commitment to sustainability (Eccles et al., 2012; Vouros et al., 2020). Nevertheless, numerous companies have faced criticism for SRs making unverified or exaggerated claims about the environmental, social, and governance (ESG) aspects of their products, services, or overall operations, potentially misleading governments, consumers, investors, and the public (Liao and Shi, 2023).

These challenges are further compounded by the limitations of traditional content analysis methods used in SR evaluations, which are often manual, time-consuming, and prone to subjectivity (Betti et al., 2018; Lindgren et al., 2021). Automated text mining techniques offer a potential solution to these challenges by enabling more scalable, consistent, and comprehensive analyses of sustainability reports. However, the application of these techniques in the context of materiality and balance is still in its infancy, and there is a lack of clarity about their effectiveness and outcomes (Lewis & Young, 2019).

1.2 Research Gap

The significant gap in the literature is twofold. First, while studies have explored materiality and balance separately, the intersection of materiality and balance remains significantly underexplored in the literature. Integrating these two dimensions to provide a holistic and accurate representation of a company's sustainability efforts, ensures that both the relevance of the reported content (materiality) and the objectivity of its presentation (balance) are addressed.

Second, there is a substantial gap in understanding how automated text mining techniques can be applied to analyze SRs and consequently address different aspects and criteria, including materiality and balance. While automated methods hold promise for overcoming the limitations of manual content analysis, their application in this context has not been fully explored. Specifically, there is a need to investigate the types of tools employed, the research objectives they can achieve, and the depth and scope of analysis these tools can provide. Understanding the current capabilities and limitations of automated text mining in the existing literature will provide a critical foundation for assessing their potential in integrating materiality and balance.

1.3 Analysis and Overall Contribution

To systematically address the identified research gaps, this thesis undertakes a two-part investigation. The first part, a systematic literature review, rigorously examines the application of text mining and natural language processing (NLP) techniques in the analysis of SR. By focusing on the current landscape, the review evaluates the methodologies employed, the specific objectives pursued, and the depth and breadth of the analyses conducted in previous studies. The goal is to assess the effectiveness and limitations of these automated techniques, particularly in the context of managing the increasing complexity and volume of SR data. This critical evaluation provides a solid foundation for advancing the application of these techniques in the field.

Then, in the second part of the investigation and by leveraging these insights, we then applied state-of-the-art techniques within a specific research context to analyze materiality and balance simultaneously. Our goal was to enhance the transparency and accuracy of reporting through more precise, automated, and rapid methods, ultimately contributing to the development of more robust and reliable sustainability assessments.

1.3.1 Use of text mining and natural language processing (NLP) techniques in analyzing sustainability reporting big data- A systematic literature review and assessment

The first research paper systematically reviews the literature on the use of text mining and NLP in the analysis of sustainability reports. This systematic literature review (SLR) contributes three key insights: First, it provides an overview of the methodologies and techniques that have been employed in the analysis of SRs. Second, it reviews the research objectives pursued by scholars utilizing TM-NLP in SR analysis. Third, it presents a critical assessment of the existing literature, revealing the depth and scope of each method applied in analyzing SRs. The findings suggest that the potential of TM-NLP to generate significant insights from SR big data remains largely unrealized, offering important directions for future research. In summary, while automated text analysis techniques show significant potential, the current application within SRs is still developing. This review highlights both the promise and the limitations of these methods, providing a comprehensive overview of the field and laying the groundwork for further advancements in the use of TM and NLP in sustainability reporting.

1.3.2 Investigating Transparency in Sustainability Reports of the Apparel Industry Using Text Mining and Natural Language Processing

The second research paper leverages advanced text mining techniques to examine the sustainability reports (SRs) of companies implicated in the Rana Plaza disaster and the broader apparel industry, analyzing both small and large datasets. By assessing materiality through theme extraction and evaluating balance by quantifying the representation of strengths, opportunities, risks, and areas for improvement, the research uncovers a significant bias toward positive reporting. The findings indicate a tendency to highlight strengths and opportunities while downplaying or positively framing risks and areas for improvement. This imbalance raises concerns about the credibility and potential greenwashing in these reports, especially when contrasted with documented misconduct in Bangladesh.

Methodologically, the study pioneers the systematic use of text analysis for SRs, offering a model for future research in theme extraction and demonstrating the value of combining sentiment analysis with theme extraction for deeper insights. Substantively, it reveals that sustainability reports in this context are skewed towards positivity, questioning the truthfulness and balance of these disclosures.

The research provides practical implications for industry practitioners, enabling companies to benchmark their sustainability reports against industry standards and improve transparency. Consulting firms can use these techniques to identify benchmarks and enhance the evaluation of corporate disclosures. By adopting these methods, stakeholders can achieve more accurate and comprehensive sustainability assessments, ultimately promoting more balanced and credible reporting practices.

1.3.3 Thesis Contribution

This thesis makes several key contributions to the field of sustainability reporting. First, it provides a comprehensive overview of the current state of automated text mining techniques as applied to SR, highlighting both the potential and the limitations of these methods. Second, by focusing on the apparel industry and the Rana Plaza disaster, it offers a critical case study that demonstrates the practical application of these techniques in a real-world context. Third, it advances the understanding of how these technologies can be used to enhance the transparency and credibility of SRs, particularly through the integrated analysis of materiality and balance, which remains a significantly underexplored area in the literature.

Together, these two studies address the identified research gaps by providing a deeper understanding of the application of automated text mining techniques in SR analysis and offering practical insights into their potential for improving the quality and reliability of sustainability reports.

Reference

Aras, G., & Crowther, D. (2008). Governance and sustainability: An investigation into the relationship between corporate governance and corporate sustainability. *Management Decision*, *46*(3), 433–448. https://doi.org/10.1108/00251740810863870

Baumüller, J., & Sopp, K. (2021). Double materiality and the shift from non-financial to European sustainability reporting: Review, outlook and implications. *Journal of Applied Accounting Research*, *23*(1), 8–28.

Betti, G., Consolandi, C., & Eccles, R. G. (2018). The relationship between investor materiality and the sustainable development goals: A methodological framework. *Sustainability*, *10*(7), 2248.

Billio, M., Costola, M., Hristova, I., Latino, C., & Pelizzon, L. (2021). Inside the ESG ratings: (Dis)agreement and performance. *Corporate Social Responsibility and Environmental Management*, *28*(5), 1426–1445. https://doi.org/10.1002/CSR.2177

Boiral, O., & Heras-Saizarbitoria, I. (2020). Sustainability reporting assurance: Creating stakeholder accountability through hyperreality? *Journal of Cleaner Production*, *243*, 118596. https://doi.org/10.1016/j.jclepro.2019.118596

Bushee, B. J., Gow, I. D., & Taylor, D. J. (n.d.). *Linguistic Complexity in Firm Disclosures: Obfuscation or Information?* Retrieved February 9, 2024, from https://onlinelibrary.wiley.com/doi/full/10.1111/1475-679X.12179?casa_token=Br0M8pMF7xkAAAAA%3A5tZj1JT-qf1L7kO_sV3P-7OvHQhAgPBnPsAw9YSWwDvAMUmSHYZ-TpyGOENCkly4g0louF6jMrgM2Q

Cerbone, D., & Maroun, W. (2020). Materiality in an integrated reporting setting: Insights using an institutional logics framework. *The British Accounting Review*, *52*(3), 100876.

Clarkson, P. M., Ponn, J., Richardson, G. D., Rudzicz, F., Tsang, A., & Wang, J. (2020). A Textual Analysis of US Corporate Social Responsibility Reports. *Abacus*, *56*(1), 1009–1051. https://doi.org/10.1111/ABAC.12182

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review*, *86*(1), 59–100.

Eccles, R. G., Krzus, M. P., Rogers, J., & Serafeim, G. (2012). The Need for Sector-Specific Materiality and Sustainability Reporting Standards. *Journal of Applied Corporate Finance*, *24*(2), 65–71. https://doi.org/10.1111/j.1745-6622.2012.00380.x

Fabrizio, K. R., & Kim, E.-H. (2019). Reluctant Disclosure and Transparency: Evidence from Environmental Disclosures. *Organization Science*, *30*(6), 1207–1231. https://doi.org/10.1287/orsc.2019.1298

Henriques, R., Gaio, C., & Costa, M. (2022). Sustainability Reporting Quality and Stakeholder Engagement Assessment: The Case of the Paper Sector at the Iberian Level. *Sustainability (Switzerland)*, *14*(21), 1–14. https://doi.org/10.3390/su142114404

Lai, A., & Stacchezzini, R. (2021). Organisational and professional challenges amid the evolution of sustainability reporting: A theoretical framework and an agenda for future research. *Meditari Accountancy Research*, *29*(3), 405–429. https://doi.org/10.1108/MEDAR-02-2021-1199

Lewis, B. W., & Young, S. (2019). Failing to report? Assessing the impact of climate risk disclosures on firm value. Accounting and Business Research, 49(5), 584-620. https://doi.org/10.1080/00014788.2019.1611730

Liao, Z., & Shi, X. (2023). Greenwashing, firm performance, and firm value: Evidence from an international sample. Energy Economics, 120, 106669. https://doi.org/10.1016/j.eneco.2023.106669

Lindgren, C., Huq, A. M., & Carling, K. (2021). Who are the intended users of CSR reports? Insights from a data-driven approach. *Sustainability, MDPI*, *13*(3), 1–20. https://doi.org/10.3390/su13031070

Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. *Critical Perspectives on Accounting*, *33*, 59–78. https://doi.org/10.1016/j.cpa.2014.10.003

Vouros, P., Nomikos, S., Halkos, G., Evangelinos, K., Sfakianaki, E., Konstandakopoulou, F., Fotiadis, S., Karagiannis, I., Skouloudis, A., & Nikolaou, I. E. (2020). Introducing fundamental accountability principles in sustainability reporting assessment: A cross-sectoral analysis from the Greek business sector. *Environmental Quality Management*, *29*(4), 33–50. https://doi.org/10.1002/tqem.21689

Chapter 1: Use of Text Mining and Natural Language Processing Techniques in Analyzing Sustainability Reports: A Systematic Literature Review and Assessment

Mahsa Mohammadrezaei

Telfer School of Management, University of Ottawa, Ottawa, Canada

PhD Thesis

University of Ottawa, Ottawa, Canada

Fall 2024

1 Introduction

Voluntary non-financial disclosures such as sustainability and corporate social responsibility (CSR) reports have received considerable attention in recent years. Companies use such disclosures to demonstrate their commitment to addressing environmental, social, and governance (ESG) issues and communicate sustainability-related information to diverse stakeholder groups (Saha & Nabareseh, 2015). Sustainability Reports (SR) play an important role in enabling investors and financial analysts to evaluate a company's present performance and make well-informed investment decisions with a purported sustainability focus (Sai et al., 2019). A variety of terms have been employed for this type of non-financial reporting, including "sustainability," "accountability," “sustainable development," "corporate social responsibility," "corporate responsibility," “non-financial disclosure,” "corporate citizenship," and “triple bottom line” reporting (M. Li & Zhao, 2021; Shahi et al., 2014). In this study, we use the term “SR”, the acronym for “sustainability report,” to refer to all these terms and any other type of similar reporting.

A sizable and growing body of evidence suggests that voluntary reporting approaches have become an expected corporate practice, effectively expanding both the quantity and length of SRs (KPMG International, 2022) - See Figure 1. However, this increase in volume has not been accompanied by an increase in the quality and precision of disclosure within them (Aras & Crowther, 2008; Billio et al., 2021; Boiral & Heras-Saizarbitoria, 2020; Bushee et al., 2018; Clark et al., 2019; Fabrizio & Kim, 2019; Kim & Lyon, 2015). Rather, the sheer volume, diversity, and complexity of the data within SRs make them challenging to evaluate, presenting considerable obstacles regarding analysis, interpretation, and comparison (Michelon et al., 2015, 2022). This has led many sustainability researchers and analysts to turn towards automated text analysis approaches such as text mining (TM) and natural language processing (NLP).

Sustainability researchers are, of course not alone in having done so. Recent advancements have made these tools more accessible and powerful, encouraging adoption across the social sciences. Some researchers and commentators have enthusiastically declared these new approaches to be transformative in nature, holding “substantial promise for automating the analysis of written and spoken language, which could transform research in economics and sociology as well as political science.” Similarly, management scholars have pointed to its capacity to transform theory development and research design within organization studies and strategic management (Barbier et al., 2022; DiMaggio, 2015; George et al., 2016; Leavitt et al., 2021; Tremblay et al., 2021). Of particular relevance to the purposes of this paper, TM-NLP has been hailed as a breakthrough

innovation capable of providing greater depth and breadth of SR analysis (Gutierrez-Bustamante & Espinosa-Leal, 2022; C. Li & Huang, 2023; Smeuninx et al., 2016; Szekely & Vom Brocke, 2017). Depth refers to the sophistication of TM-NLP techniques used to extract detailed and nuanced insights from sustainability reports, while breadth pertains to the ability of these methods to systematically analyze large-scale datasets across multiple reports, industries, and time periods. Combining high depth and breadth enables a comprehensive and insightful understanding and assessment of SR quality.

Automated text analysis undoubtedly holds great promise for dealing with the growing volume and complexity of SRs, thereby opening up new avenues of research and insight, and provide robust and actionable insights from SR data. For example, NLP, a specialized branch of Machine Learning (ML) applications is dedicated to working with human language, particularly textual data. NLP encompasses a broad range of techniques aimed not only at extracting information from text (i.e., text mining) but also at comprehending, interpreting, and even generating human-like language. These techniques span various tasks such as machine translation, text summarization, and sentiment analysis. NLP can be operationalized through supervised , unsupervised ML algorithms or through a combination of both to achieve these tasks effectively. Supervised algorithms are trained on a labeled dataset, where each data point is associated with a known output or target value. The algorithm learns to map the input data to the output data, and it can then be used to make predictions on new, unseen data. Unsupervised algorithms, on the other hand, do not require labeled data. Instead, they learn to identify patterns and structures in the data itself which can be useful for tasks such as clustering and dimensionality reduction (James et al., 2021; Nadkarni et al., 2011). While it has the ability to work with unlabeled data, which is more readily available and does not require lots of user input in the front end, interpreting the results of unsupervised ML in the backend can be more challenging given that the results may be complex or nuanced. and require the user to interpret (Nadkarni et al., 2011).

There is, however, considerable uncertainty and a lack of understanding regarding the application and outcomes of TM-NLP techniques (Lewis & Young, 2019; Miner et al., 2012; Nishant et al., 2020). We focus on clarifying two main aspects. First, given the novelty of this approach, we lack a clear overview and mapping of the types of tools employed and the research objectives that have been achieved with these tools. Put simply, there is a lack of clarity regarding the means and ends of automated text analysis of SRs. This is partly a reflection of the fact that much of what has been written either offers generalizations regarding the potential of TM-NLP, or alternatively, provides fine-grained technical details about the research techniques themselves without saying much about

what can be achieved with these tools and approaches. We begin by clarifying what is possible in terms of research objectives and what tools are available for reaching these objectives.

Second, there has been no comprehensive assessment of the progress made in applying automated text analysis tools to sustainability reports (SRs). In other words, the extent to which this technology's capabilities have been realized remains unclear. To address this gap, we offer an overview of the current state of these tools and techniques as applied to SR analysis. Specifically, we conduct a systematic literature review (SLR) focusing on studies that utilize one or more text mining and natural language processing (TM-NLP) methods to analyze SRs. We focus on assessing research depth and breadth.

Based on this overview of the state-of-the-art techniques and our systematic review of the research methods and objectives, we provide an assessment of the techniques employed for the automated text analysis of SRs. We find that although there has been notable advancement in this area, our review identifies several key challenges related to research depth, breadth, rigour, and transparency. These challenges suggest that the full capabilities of text mining and natural language processing (TM-NLP) to extract valuable insights from the vast amounts of SR data have yet to be fully realized.

This literature review addresses two key research questions: RQ1: What are the different text mining methods employed to analyze SRs? and RQ2: What are the objectives of researchers using these methods? By answering these questions, the review provides a clear understanding of the tools and techniques applied in SR analysis and the goals driving this research. It highlights the untapped potential of TM-NLP for enhancing SR quality and identifies critical issues for future research. Addressing these gaps will allow for better use of TM-NLP tools to improve SR analysis, offering valuable insights for organizations, policymakers, and stakeholders. The findings suggest that TM-NLP holds significant promise for automating the analysis of SRs, which can enhance the quality of sustainability reporting and support research in other fields, such as economics, sociology, and political science.

The remainder of this paper is organized in the following manner. The next section provides some background on automated text analysis techniques for SR research that is germane to our review. This includes a description of the main issues concerning SR big data and the main benefits touted by proponents of automated text analysis techniques. Section 3 then explains our SLR methodology. We describe the search strategy that was used and the data extraction process. Section 4 presents the detailed findings across the two research questions posed in this review and elaborated in Section 3.

Section 5 provides a discussion of the findings, critical reflections on the reviewed literature, and outlines suggestions for future research.

2 Automated Text Analysis Technique for Sustainability Report Research

While automated text analysis has received prominent attention, there is a lack of clarity and conflation regarding some of its key approaches, as well as the capabilities and benefits it brings to the analysis of SR big data. Therefore, before presenting our study's methodology and findings, we define and explain some key concepts. We then situate our review in both current literature and practice by providing a high-level summary of what are generally considered to be the main challenges of manual analysis of sustainability data and the reasons for researchers to adopt automated text analysis techniques. It should be noted that this is neither an exhaustive explanation of the concepts underpinning TM-NLP nor a full review of the literature examining the challenges of manual text analysis versus the benefits of automated text analysis.

2.1 Key terms and concepts

NLP is closely related to text mining but goes one step further to focus on understanding, interpreting, and generating human-like language. This can include tasks such as machine translation, text summarization, and sentiment analysis (Antons et al., 2020). Some aspects of NLP can be understood as text mining, whereas some are more sophisticated. NLP represents a specialized branch of Machine Learning (ML) applications dedicated to working with human language, particularly textual data. NLP encompasses a broad range of techniques aimed not only at extracting information from text (i.e., text mining) but also at comprehending, interpreting, and even generating human-like language. These techniques span various tasks, such as machine translation, text summarization, and sentiment analysis.

In NLP applications, both supervised and unsupervised ML algorithms are utilized to achieve these tasks effectively. Where required for understanding and accuracy, we specify whether these are text mining or NLP. Otherwise, we use the acronym TM-NLP when referring to the general category of automated text analysis methods as a means of simplifying analysis and discussion (Nadkarni et al., 2011).

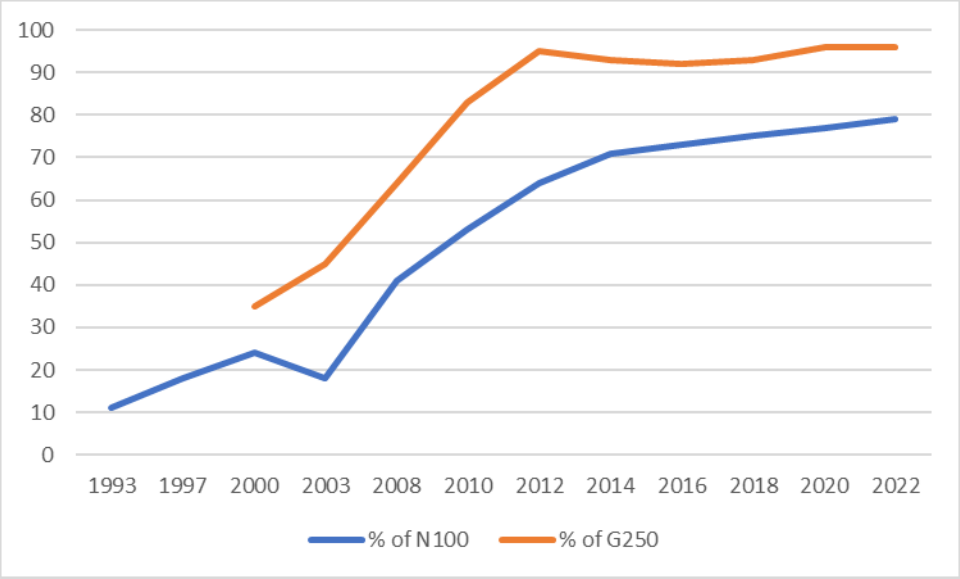
Another set of terms that require definition is "supervised learning" and "unsupervised learning." Supervised and unsupervised learning are general categories of ML methods that can be applied to a wide variety of data types, including text, images, audio, and video. Supervised ML algorithms are trained on a labelled dataset, where each data point is associated with a known output or target value.

The algorithm learns to map the input data to the output data, and it can then be used to make predictions on new, unseen data. As each example has a known output or target value, the model can learn the relationship between the input data and the output data, which makes it easier to interpret the results. Unsupervised ML algorithms, on the other hand, do not require labelled data. Instead, they learn to identify patterns and structures in the data itself, which can be useful for tasks such as clustering and dimensionality reduction (James et al., 2021; Nadkarni et al., 2011). For instance, in topic modelling, which we refer to below, an unsupervised ML algorithm might identify prevalent topics in a collection of documents. As such, there are both benefits and drawbacks to working with unsupervised ML methods. While it has the ability to work with unlabeled data, which is more readily available and does not require lots of user input in the front end, interpreting the results of unsupervised ML in the backend can be more challenging, given that the results may be complex or nuanced. And require the user to interpret (Nadkarni et al., 2011).

2.2 The case for TM-NLP methods in analyzing SR

In making the case for TM-NLP methods in analyzing SRs, three central issues are frequently mentioned – growth in volume, lack of standardization, and poor informational quality. First, the volume of qualitative data disclosed is growing remarkably due to the surge in the number of organizations issuing SRs (see Figure 1) and the steady increase in the content contained within the reports themselves. The term “big data” has become increasingly used to describe this growing body of data (Kang & Kim, 2022; Li & Huang, 2023; Wanner & Janiesch, 2019). Second, the data provided in the SRs are non-standard in nature and format. Reports include images, numbers, and texts, which are presented in a variety of ways (e.g., narrative descriptions, tables, charts, graphics, etc.), as well as an increasing diversity of languages. This is due in part to a proliferation of ESG reporting formats, standards, and frameworks for disclosing non-financial data that have evolved in recent decades, resulting in a confusing array of foci and what is considered “material,” resulting in notable difficulties in comparing this growing volume of data structures (Eccles et al., 2011; Kozlowski et al., 2015; Luo et al., 2015; Roca & Searcy, 2012). Third, the reported data is often of poor quality or a misleading nature, what is often referred to as “greenwash” .”is relates to a number of observed and documented problems, including issues or actions discussed that may not be relevant or “material” to the industry or sector of activity; information provided is generic, vague, or devoid of meaningful content; information is misleading or presented in a manner to purposefully obfuscate; omits important and relevant information (Marquis et al., 2016; Yu et al., 2020).

Figure 1: SR rates among the world’s largest companies



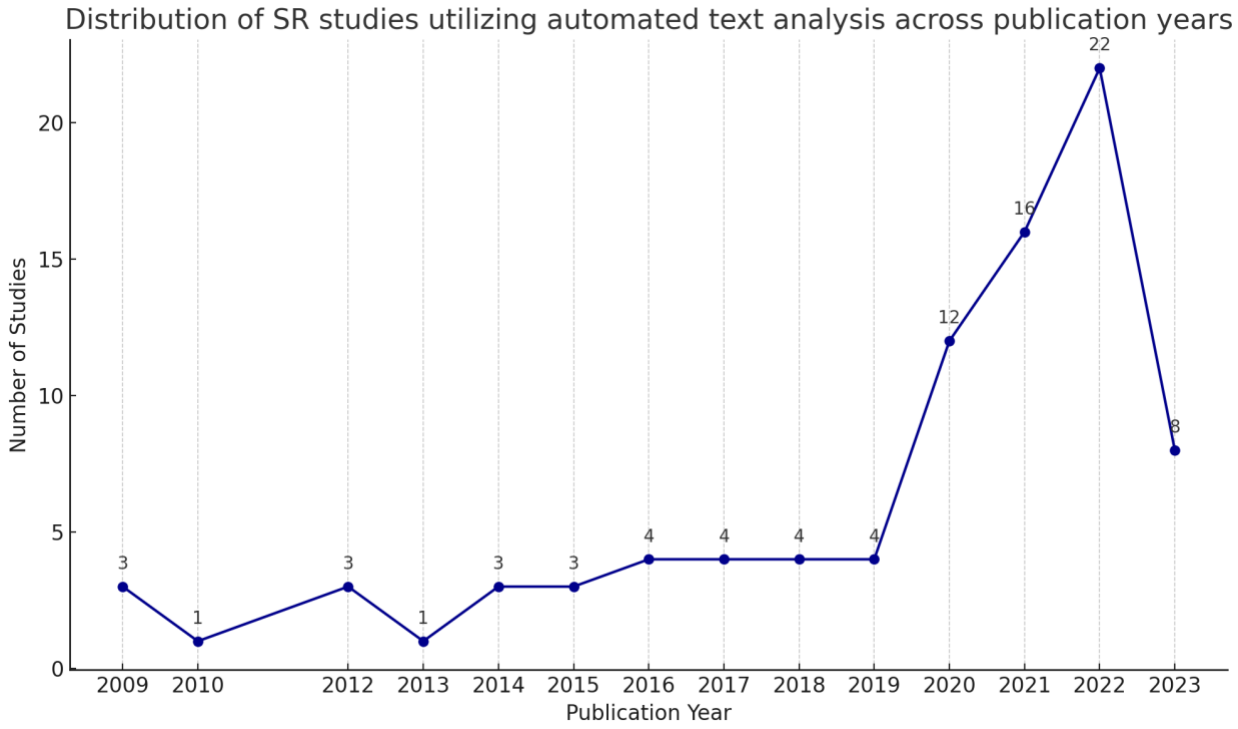
G250 refers to the world's biggest 250 companies by revenue; N100 refers to the top 100 businesses in each of 58 countries (total of 5,800 companies).

Source: KPMG Survey of SR2022, KPMG International, September 2022

https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2023/04/big-shifts-small-steps.pdf

As a result of these issues, the manual analysis of SRs and the ability to extract useful information for decision-making have become prohibitively costly and time-consuming (Shahi et al., 2014). For analysts and researchers, the systematic collection and analysis of this data has become an immense undertaking, often entailing the cataloging and coding of tens of thousands of pages of documents (Fiandrino & Tonelli, 2021). This has led to an interest in, and demand for, automated text analysis solutions for examining and scoring SRs (see Figure 2) (Harymawan et al., 2020; Shahi et al., 2014; A. Zhou, 2021) , with ESG analysts and academics increasingly employing various TM and NLP approaches to analyze company disclosures where qualitative data is predominant (e.g., K-2 filing in the US, auditor reports, and SRs) (Aureli et al., 2016; Harymawan et al., 2020; Kiriu & Nozaki, 2020a; Liew et al., 2014; Mohan et al., 2016; Reuter et al., 2014; Tremblay et al., 2015).

Figure 2: Distribution of SR studies utilizing automated text analysis across publication years



When compared to manual approaches, this automated text analysis form of content analysis offers several advantages. First, it increases efficiency and reduces the time required for analysis. Using algorithms, researchers can process and analyze large volumes of data much quicker, allowing for more comprehensive and extensive investigations (Deng et al., 2017; Fiandrino & Tonelli, 2021; Tremblay et al., 2015). Second, automated text analysis enhances reliability by minimizing human error, subjectivity, and bias. This helps mitigate errors and discrepancies that may inadvertently arise as a result of human interpretation or subjective judgment during the manual coding process. Automated text analysis eliminates these issues by employing standardized algorithms that adhere to predefined coding rules, resulting in a more objective, accurate, and consistent analysis across the entire dataset (Q. Deng et al., 2017a; Kiriu & Nozaki, 2020a). Third, automated text analysis is cost-effective. It eliminates the need for a large workforce dedicated to manual coding and reduces the need for error correction and data verification, thereby substantially reducing labour costs. Additionally, the use of automated tools and software reduces the expenses associated with manual error correction and data verification (Kiriu & Nozaki, 2020a; M. Li & Zhao, 2021).

Overall, these three benefits suggest that TM and NLP have provided the possibility of analyzing substantial volumes of sustainability data and the ability to generate notable new insights. We refer to these as research depth and research breadth. Research depth refers to the sophistication of TM-NLP techniques, which encompasses the complexity of the algorithms used, the extent of contextual

understanding they provide, and their ability to produce nuanced insights(Blei et al., 2003a; Jurafsky, 2000; Manning et al., 2008). The level of research depth is critical because it determines how effectively these tools can go beyond surface-level analysis to uncover complex details within SRs (Feldman & Sanger, 2007; Miner et al., 2012). For instance, sophisticated algorithms can parse through dense corporate language, identify nuanced differences in how sustainability practices are reported, and detect the sentiment behind the narratives presented. This level of detail is essential for understanding the true intent behind a company's sustainability claims, discerning whether the language used reflects genuine commitment or is more aligned with "greenwashing" (Delmas & Burbano, 2011; Torelli et al., 2020). Moreover, higher research depth enables the identification of subtle patterns, such as shifts in a company's sustainability focus over time or the emergence of new themes in response to changing regulatory or market pressures. It can also reveal underlying themes that might not be immediately apparent, such as the integration of sustainability into corporate governance or the impact of sustainability initiatives on financial performance.

On the other hand, research breadth pertains to the range and volume of data that these TM-NLP techniques can systematically analyze. This aspect is particularly prominent in the context of "big data," where the challenge lies not only in processing vast amounts of text but also in doing so efficiently across multiple SRs, industries, and periods. Effective research breadth means that these tools can handle large-scale datasets, analyzing thousands of reports to identify commonalities and differences across industries or regions (Aggarwal & Zhai, 2012; Miller & Mork, 2013). This capability is Imperative for generating insights that are not limited to a single company or sector but are instead applicable on a broader scale, offering a more comprehensive understanding of sustainability trends and practices. For instance, by analyzing reports from different industries, researchers can identify sector-specific sustainability challenges and strategies, which can then inform industry-wide best practices. Additionally, research breadth allows for the systematic analysis of changes over time, helping to track the evolution of sustainability reporting practices and the impact of global sustainability initiatives, such as the United Nations' Sustainable Development Goals (SDGs).

In essence, combining high research depth with substantial research breadth enables a more robust and insightful analysis of SRs. While research depth ensures that the analysis is detailed and nuanced, uncovering the complex dynamics within sustainability narratives, research breadth ensures that these insights are drawn from a comprehensive dataset, enhancing their generalizability and relevance. Together, these dimensions of TM-NLP contribute to a more holistic understanding of

corporate sustainability, allowing stakeholders to make more informed decisions based on a thorough analysis of sustainability reports.

3 Research Method

To investigate the present state of research on the application of TM techniques in analyzing SR, we conduct a systematic literature review of the relevant literature. To build our SLR methodology, we follow Kitchenham et al. (2009) .

3.1 Search strategy

To ensure the identification of relevant studies, it is imperative to employ appropriate search concepts, phrases, and words. In this study, we utilized a comprehensive search query in Scopus, a widely recognized and authoritative digital library in the academic community. The search was implemented as follows:

("NFR*" OR "non-financial report*" OR "non financial report*" OR "Sustainability report*" or "CSR report*" or "corporate social responsibility report*" or "environment* disclosure" or "esg report*") AND ("NLP" or "Natural language processing," or "text mining" or "ml" or "machine learning" or "content analysis software" or "text* analysis" or "topic model*").

The inclusion of terms such as "NLP," "natural language processing," "text mining," "ml," and "machine learning", "content analysis software", "text* analysis" and "topic model*" ensure that studies utilizing various techniques were captured. Additionally, we incorporated terms like "Sustainability report*," "CSR report*," "corporate social responsibility report*," "NFR*," "esg report*," and "environment* disclosure" to specifically focus on articles that addressed SRs using the various text mining techniques. By carefully selecting and combining these search concepts, phrases, and words, we aimed to identify the most relevant studies within the literature in the focus of this review.

For our SLR, we chose to focus our search exclusively on the Scopus digital library, considering only peer-reviewed journal articles and conference papers, while excluding book chapters and other publication formats. Conference papers were selected because TM-NLP methods are typically developed within the computer science domain, where conference papers are a more common and timely form of disseminating new methods and applications compared to other publication types. Scopus was chosen due to its comprehensive coverage across various disciplines, including management, sustainability, and computer science. This selection aimed to ensure a representative sample of relevant, reliable, and scholarly research. The search was conducted on July 25th, 2023,

without imposing any restrictions on the publishing year, to encompass a broad range of literature and to also include the most current and seminal research in the field. By considering articles from different time periods, we aimed to obtain a comprehensive understanding of the existing knowledge on the topic.

In the first step of our SLR, a thorough screening process was conducted to filter out articles based on predefined inclusion and exclusion criteria, ensuring replicability and rigour in the review. This rigorous screening involved a careful examination of titles, abstracts, and keywords to identify relevant papers aligned with the review objectives. In the next step, a full-text screening of the remaining papers was done. This comprehensive evaluation allowed for a more in-depth assessment of each paper's content and relevance to this review. Any paper that was deemed irrelevant or did not meet the predetermined criteria was excluded from further consideration, maintaining the integrity of the review process. Studies that did not analyze the SRs' contents, studies that did not use any TM or NLP techniques in their methodology or lacked a clear identification of TM methodology, studies that were not peer-reviewed or were in the form of posters, proceedings, short papers, abstracts, book chapters, or were review and survey studies, and studies that were not in English were excluded from the final selection. Based on the initial search, we obtained 241 results. Applying the exclusion criteria, a final selection of 88 papers was made, comprising the most pertinent and valuable contributions that met the quality and inclusion criteria for our systematic literature review.

3.2 Data Extraction

The data extraction stage involved collecting relevant data and information from the 88 selected publications. This included the publication outlet, key research themes and topics, the objective for analyzing SRs, the utilized TM-NLP methods and techniques, the number of sustainability reports (SRS) analyzed (data scope), and the time span of SRs being included in those studies. The resulting data from this process served as the foundation for the subsequent analysis.

4 Findings

4.1 Overview of NLP-TM methods used in analyzing SRs

Our analysis of the reviewed papers reveals that certain NLP-TM methods have been employed for the automated text analysis of SRs. These methods, ranked by frequency of use, include topic modelling, clustering, word frequency analysis, and semantics-based approaches. Additionally, several studies employed proprietary software tools that streamline the application of these methods.

This section provides an overview of each of these methods, along with relevant techniques and an analysis of their computational depth and breath.

4.1.1 Topic-modeling method and techniques

Topic modelling is a critical NLP-TM method used to uncover latent topics within a collection of documents, such as SRs, by analyzing patterns in word usage and grouping them into thematic clusters. Among the methods identified, Latent Dirichlet Allocation (LDA), Latent Semantic Analysis (LSA), K-Nearest Neighbors (KNN), and BERT were most frequently employed. Each technique has unique strengths and weaknesses.

- *LDA:* An unsupervised method, LDA is particularly adept at identifying a mixture of topics within documents and associating specific words with each topic. This Bayesian probabilistic approach excels in handling large datasets, making it a popular choice for studies requiring deep thematic analysis of SRs. LDA's ability to discern topics such as environmental impact and corporate governance across numerous SRs has proven invaluable in studies requiring detailed thematic exploration (Benites-Lazaro et al., 2018; Hadro et al., 2022). Its application on datasets ranging from 27 to 9,514 SRs showcases its scalability and robustness in diverse contexts (Szekely & Brocke, 2017; Calabrese et al., 2023).
- *LSA:* Utilizing Singular Value Decomposition (SVD), LSA reduces the dimensionality of text data, capturing relationships between words and documents in a lower-dimensional space. This method is effective for identifying latent structures in SRs, though it may miss nuanced contextual relationships compared to LDA. LSA has been successfully applied to medium-sized datasets, often between 100 to 1,000 SRs, providing insights into the co-occurrence of terms like "renewable energy" and "carbon footprint" (Kountouri et al., 2019; Pan, 2016).
- *KNN:* Typically used for classification and clustering, KNN's simplicity allows it to identify document similarities based on feature vectors, making it useful for preliminary thematic grouping in SRs. Its application is generally limited to medium datasets due to its computational inefficiency with larger corpora (Raghupathi et al., 2020).
- *BERT:* A transformer-based model developed by Google, BERT excels at capturing context by considering both preceding and following words in a text. Although BERT is computationally expensive, it has demonstrated superior performance in detailed text classification and sentiment analysis within SRs, particularly in studies involving nuanced sustainability disclosures (Devlin et al., 2018; Wang et al., 2020). Its ability to handle

complex data with medium-sized datasets highlights its capacity for deeper semantic analysis, despite being underutilized due to resource constraints.

4.1.2 Supervised Methods and Classifiers

Supervised methods and classifiers offer high accuracy in categorizing SRs based on labelled datasets. These methods involve training algorithms on a labelled dataset, allowing the model to learn patterns and relationships that can be applied to new, unseen data. These methods include various ML algorithms that excel in text categorization and sentiment analysis:

- *Neural Networks:* Modeled after the human brain, these algorithms are capable of recognizing complex patterns in text, making them suitable for complex tasks like SR categorization and pattern recognition. However, they require extensive computational resources and well-labelled data (Gutierrez-Bustamante & Espinosa-Leal, 2022).
- *Random Forest and XGBoost:* These ensemble learning techniques combine multiple decision trees to enhance prediction accuracy and mitigate overfitting. Their application in sentiment analysis and readability assessments has shown robustness and high accuracy, particularly in large-scale SR datasets (Clarkson et al., 2020; D'Amato et al., 2021).
- *Naive Bayes:* Despite its simplicity, this probabilistic classifier remains effective for text classification tasks, serving as a reliable baseline for more complex models (Shahi et al., 2014).

4.1.3 Clustering and Unsupervised Techniques

Clustering and unsupervised techniques are integral to identifying patterns in SRs by grouping documents based on content similarity:

- *K-means Clustering:* A widely-used algorithm for partitioning data into K clusters, K-means has been instrumental in revealing thematic structures within SRs. Its simplicity and efficiency make it suitable for large datasets, although it requires the pre-specification of cluster numbers and is sensitive to initial conditions (Liu et al., 2017).
- *Hierarchical Clustering:* By creating a tree-like structure of clusters, hierarchical clustering provides a visual understanding of relationships within data, making it useful for SRs with complex thematic hierarchies (Liu et al., 2017).

- *DBSCAN:* This density-based algorithm is effective for identifying clusters of arbitrary shapes and handling noise, making it valuable in exploratory analyses of SRs where data may be unevenly distributed (Liu et al., 2017).
- *Word2Vec:* A neural network-based technique that generates vector representations of words, Word2Vec captures semantic relationships within text, facilitating the identification of thematic clusters. It is particularly effective for large-scale text analysis, offering deep semantic insights into SRs (Bodendorf et al., 2022).

4.1.4 Word frequency analysis method and techniques

Word frequency analysis provides a foundational approach to understanding the content of SRs by identifying the most commonly occurring terms:

- *TF-IDF:* This method highlights important terms within a document relative to a corpus by balancing term frequency with inverse document frequency. TF-IDF has been widely used for evaluating the importance of terms in SRs and is particularly useful for preliminary thematic analysis (Raghupathi et al., 2020).
- *Keyword Matching:* A straightforward technique that involves searching for predefined keywords within SRs, keyword matching is effective for compliance assessments but may miss nuanced language that falls outside the predefined terms (Moreno & Caminero, 2022).

4.1.5 Semantic methods and techniques

Semantics-based methods explore the linguistic aspects of SRs, analyzing sentiment, readability, and linguistic complexity:

- *Sentiment Analysis:* Techniques like sentiment analysis assess the emotional tone of SRs, classifying content as positive, negative, or neutral. These methods have been applied to understand public perception and the overall sentiment conveyed in sustainability disclosures (Harymawan et al., 2020; Sai et al., 2019).
- *Readability Measures:* Tools like the Fog Index measure the complexity of SRs by analyzing sentence length and word difficulty. These measures provide insights into how accessible and understandable the reports are to their intended audience (Smeuninx et al., 2016).

4.1.6 Proprietary software-based methods and techniques

Several studies employed proprietary software tools for SR analysis. These tools, while not as advanced as the latest NLP approaches, offer user-friendly interfaces and a range of techniques for extracting insights from unstructured text data:

- *Leximancer:* This tool provides visual concept mapping, making it useful for thematic analysis and identifying relationships between different topics within SRs (Vaio et al., 2022).
- *RapidMiner:* Known for its flexibility and comprehensive analytics capabilities, RapidMiner enables automated analyses and is particularly effective in large-scale text mining tasks (Wang et al., 2020).
- *Provalis WordStat:* Offering keyword extraction, content analysis, and text visualization, WordStat is widely used for social sciences research, providing detailed qualitative insights into SRs (Tóth et al., 2021).
- *DICTION:* This software specializes in sentiment and readability analysis, providing metrics on emotional content and text complexity. It has been employed in studies examining the tone and effectiveness of SR communication strategies (Feng & Gao, 2020).

4.2 Analysis of TM-NLP Methods from a Computer Science Perspective

In this section, we categorize the various methods employed in SR analysis by their computational complexity (depth) and data handling capacity (breadth), providing a framework for selecting the most appropriate techniques based on specific analytical needs.

Depth Analysis

- *High-Depth Methods*: Techniques such as BERT, Random Forest, XGBoost, Neural Networks, LDA, and Word2Vec require significant computational resources and excel in handling complex relationships within large datasets. These methods are ideal for in-depth text categorization, sentiment analysis, and topic modelling.
- *Medium to advance Depth Methods*: LSA, K-means, Hierarchical Clustering, DBSCAN, and association-based TM strike a balance between computational complexity and analytical capability. They are effective for pattern identification and thematic analysis but may not fully capture deeper semantic relationships.
- *Low Depth (Surface level) Methods*: Methods such as Naive Bayes, keyword matching, word frequency analysis, and readability measures are computationally simple and fast. They are

suitable for preliminary analyses and large datasets but lack the ability to delve into deeper semantic and contextual nuances.

Breath Analysis

In terms of data handling capacity, methods can be categorized as follows:

- *Large Data Capacity Methods:* These methods are highly suitable for large-scale text analysis, capable of handling thousands to millions of records efficiently. They offer robust performance and scalability, making them ideal for analyzing extensive datasets.
- *Medium to small Data Capacity Methods:* These methods are effective for medium-sized datasets, handling hundreds to thousands of records. They offer a balance between managing moderately large datasets and providing detailed insights.

This detailed analysis provides a strong foundation for selecting appropriate methods based on the specific requirements of depth and breadth in SR analysis. By understanding the strengths and limitations of each method, researchers can make informed decisions to achieve accurate and meaningful insights from sustainability reports.

The following table elaborates on the techniques, categorizing them into advanced, medium to advanced, surface-level, and software-based classifications. It offers a comprehensive analysis of their inherent strengths, weaknesses, depth, and breadth from a computer science perspective. The references cited either utilized these methods or provided in-depth explanations, enabling us to capture these inherent capabilities and create the following table. We have provided this table to give researchers a clearer understanding of the inherent strengths of the methods extracted from the analyzed papers, offering valuable insights into their capabilities.

Table 1: Comprehensive Table of Methods from a Computer Science Perspective

Category	Method/Technique	Depth	Breadth	Strengths	Weaknesses	Reference
Advanced Methods	BERT	High	Medium (Hundreds to thousands)	Captures nuanced semantic relationships, high accuracy	Requires extensive computational resources, large training datasets	(Devlin et al., 2018)
	Random Forest	High	Large (Thousands to millions)	Robust, handles high-dimensional data, reduces overfitting	Requires substantial computational resources, complex model tuning	(Liaw & Wiener, 2002)

Category	Method/Technique	Depth	Breadth	Strengths	Weaknesses	Reference
	XGBoost	High	Large (Thousands to millions)	High accuracy, efficient handling of missing data, flexible	Computationally expensive, requires careful parameter tuning	(T. Chen & Guestrin, 2016)
	Neural Networks	High	Large (Thousands to millions)	High accuracy, handles complex relationships, adaptable	Requires substantial computational resources, extensive training time	(Goodfellow et al., 2016)
	LDA	High	Large (Thousands to millions)	Uncovers latent topics, handles large datasets	Computationally intensive,	(Blei et al., 2003b)
	Word2Vec	High	Large (Thousands to millions)	Captures semantic relationships, useful for clustering and theme extraction	Requires substantial computational resources, large training datasets	(Mikolov et al., 2013)
Medium to Advanced Methods	LSA	Moderate	Medium (Hundreds to thousands)	Reduces dimensionality, reveals hidden structures	May miss contextual nuances, less effective with small datasets	(Kountouri et al., 2019)
	K-means Clustering	Moderate	Large (Thousands to millions)	Simple, efficient for large datasets, easy to implement	Requires specifying number of clusters, sensitive to initial conditions	(MacQueen, 1967)
	Hierarchical Clustering	High	Medium (Hundreds to thousands)	Provides detailed cluster hierarchy, easy to interpret	Computationally intensive for large datasets, memory intensive	(Johnson, 1967)
	DBSCAN	Moderate	Medium (Hundreds to thousands)	Identifies clusters of arbitrary shape, handles noise	Requires parameter tuning, less effective with varying densities	(Ester et al., 1996)
	Association-Based Text Mining	Moderate	Medium (Hundreds to thousands)	Reveals relationships between concepts, useful for exploratory analysis	May miss deeper semantic relationships, dependent on term frequency	(Agrawal & Srikant, 1994)
	Factor Analysis	High	Medium (Hundreds to thousands)	Identifies underlying factors, reduces dimensionality	Complex interpretation,	(Bartholomew et al., 2011)

Category	Method/Technique	Depth	Breadth	Strengths	Weaknesses	Reference
					requires careful preprocessing	
	Cosine Similarity	Moderate	Medium (Hundreds to thousands)	Measures similarity between documents, easy to implement	Sensitive to document length, may not capture semantic meaning	(Hummel et al., 2022)
	KNN Clustering	Moderate	Medium (Hundreds to thousands)	Simple, easy to implement, non-parametric	Computationally expensive for large datasets, sensitive to noise and irrelevant features	(Cover & Hart, 1967)
Surface Level Methods	Naive Bayes	Low	Large (Thousands to millions)	Fast, simple, effective for text classification	Assumes feature independence, less effective with highly correlated data	(Manning et al., 2008)
	Keyword Matching	Low	Large (Thousands to millions)	Fast, easy to implement	Limited to predefined keywords, lacks contextual understanding	(Manning et al., 2008)
	TF-IDF	Moderate	Medium (Hundreds to thousands)	Highlights important terms, simple implementation	Surface-level analysis, lacks contextual depth	(Manning et al., 2008)
	Word Frequency Analysis	Low	Large (Thousands to millions)	Provides initial insights, easy to implement	Lacks depth, misses semantic relationships	(Manning et al., 2008)
	Sentiment Analysis	Moderate	Medium (Hundreds to thousands)	Analyzes tone and emotion, useful for understanding public perception	May oversimplify sentiment, lacks nuanced understanding	(Pang & Lee, 2008)
	Readability Measures	Low	Large (Thousands to millions)	Quantifies text complexity, easy to calculate	Limited depth, may miss linguistic nuances	(Collins-Thompson, 2014)
Software-Based Methods	Leximancer	Moderate	Medium (Hundreds to thousands)	Provides visual representation, user-friendly	Limited customization, may not capture deep semantics	(Thomas, 2014)

Category	Method/Technique	Depth	Breadth	Strengths	Weaknesses	Reference
	RapidMiner	High	Large (Thousands to millions)	Comprehensive analytics, flexible	Requires learning curve, computationally intensive	(Kalra & Aggarwal, 2017)
	Provalis WordStat	Moderate	Medium (Hundreds to thousands)	Integrated tools, user-friendly	Limited to available features, may require manual adjustments	(Son, 2005)
	DICTION	Moderate	Medium (Hundreds to thousands)	Efficient and automated, suitable for thematic analysis	Limited by predefined dictionaries, may lack depth	(Feng & Gao, 2020)

4.3 Research Objectives and Methodological Frequency/Intersections in Analyzing Sustainability Reports (SRs)

Our analysis of 88 selected publications reveals three primary research objectives in the use of NLP-TM methods for SR analysis: Identification of SR Themes and Patterns, Analysis of SR Compliance with Existing Frameworks, and Sentiment and Readability Assessment.

- *Objective 1:* Identification of SR Themes and Patterns: This objective focuses on uncovering the underlying themes and patterns within SRs. Advanced methods like LDA and BERT were predominantly used, offering deep insights into latent topics within SRs. These techniques facilitated the extraction of complex themes and patterns from large datasets. In many cases, these advanced methods were complemented by surface-level techniques like TF-IDF, which provided initial thematic exploration and supported the deeper analysis conducted by more sophisticated models.
- *Objective 2:* Analysis of SR Compliance with Existing Frameworks: The goal here is to evaluate the extent to which sustainability reports align with established non-financial reporting standards such as GRI, SASB, and SDGs. A range of methods, from surface-level techniques to more advanced unsupervised methods, were used to assess compliance with predefined frameworks such as GRI, SASB, and SDGs. While surface-level techniques and software-based tools were frequently employed for preliminary compliance assessments, unsupervised methods like LDA and clustering were also applied in some studies to explore

deeper compliance patterns. However, the use of these advanced methods was less common, suggesting opportunities for methodological enhancement in this domain.

- *Objective 3:* Sentiment and Readability Assessment: This objective aims to analyze the emotional tone and linguistic complexity of SRs, which can reflect a company's performance and the challenges it faces. Sentiment analysis and readability measures were commonly used to evaluate the tone and complexity of SRs. Techniques such as BERT were particularly effective in capturing nuanced sentiment and contextual relationships within the text, although their application was often restricted to smaller datasets due to high computational demands. Other methods, including dictionary-based approaches and simpler sentiment analysis techniques, were employed to analyze larger datasets, providing broad overviews of sentiment and readability while being less resource-intensive.

Figure 3 illustrates the number of methods found in the papers, categorized by their level of complexity: advanced, medium to advanced, surface level, and software-based methods. This categorization highlights the frequency of each method within these categories. Medium to advanced techniques, including LSA, Cosine Similarity, and K-means Clustering, are the most prevalent, indicating their balanced approach to computational complexity and analytical capability.

Figure 3 The Distribution of TM-NLP Methods by Complexity Level

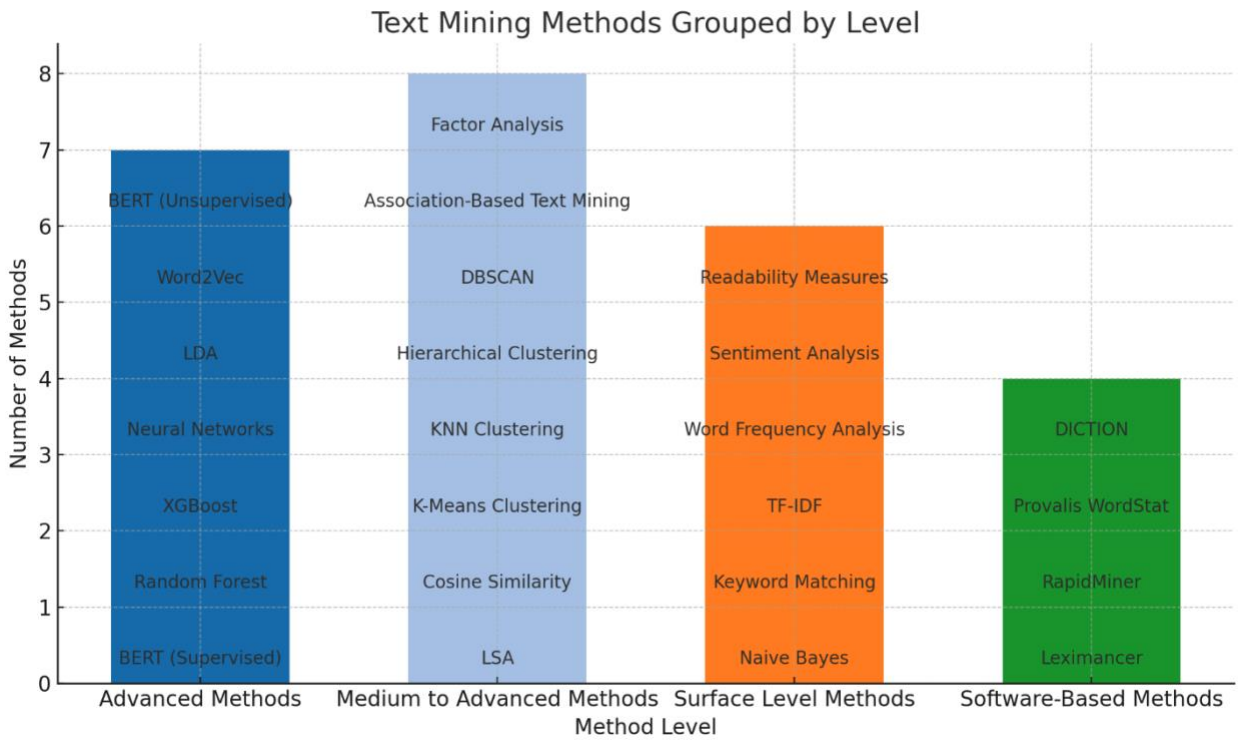


Figure 4 provides insight into the frequency of each method used across the three primary objectives identified in the review. The "Identification of SR Themes and Patterns" objective is the most represented, with LDA being the most frequently used method due to its ability to uncover latent topics within large datasets. This objective also utilizes a mix of advanced and surface-level techniques, highlighting a balanced approach that combines deep semantic analysis with straightforward frequency-based methods.

For the Analysis of SR Compliance with Existing Frameworks, the reliance on surface-level and software-based techniques indicates a preference for methods that offer ease of implementation and scalability, albeit at the cost of depth. The limited use of advanced techniques in this objective suggests an opportunity for more sophisticated approaches in future research.

The Sentiment and Readability Assessment objective, though less represented, employs both advanced methods like BERT and simpler readability measures. This reflects the need for both deep contextual understanding and straightforward assessments in evaluating the sentiment and readability of SRs.

Figure 4 Distribution of methods for each objective

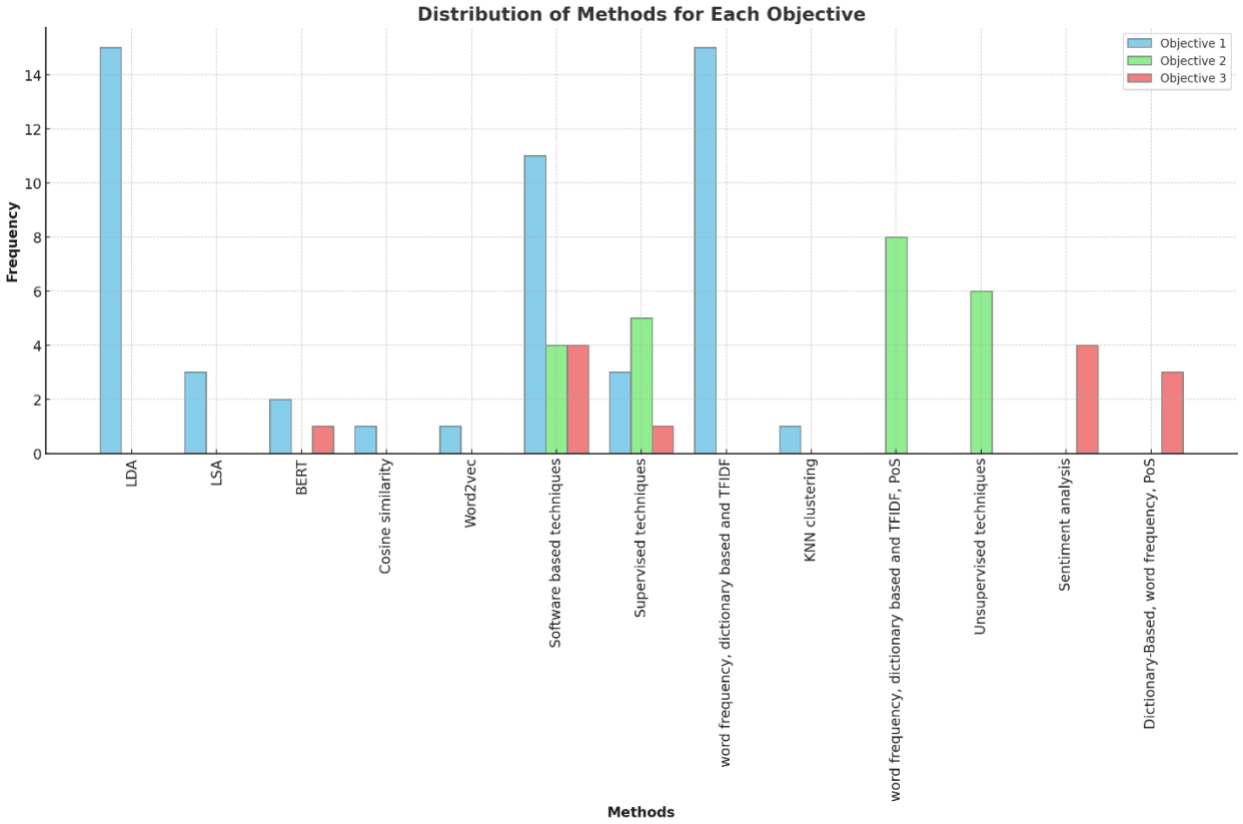
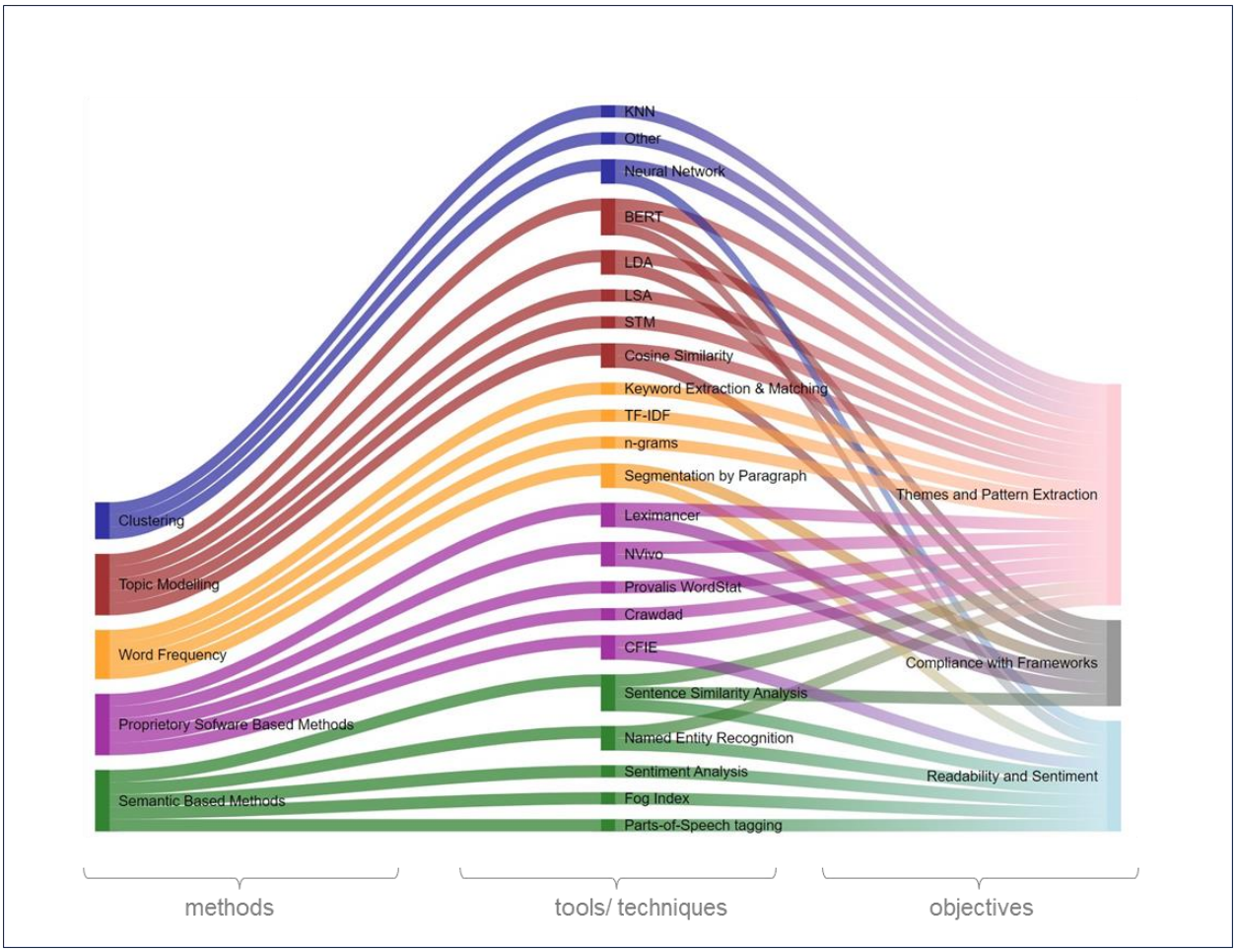


Figure 5 visualizes the intersection between research objectives, NLP methods, and the specific tools/techniques used. This mapping highlights the versatility and effectiveness of certain methods, such as BERT and LDA, which are applied across multiple objectives. The figure demonstrates the adaptability of these techniques in various analytical contexts and offers a comprehensive view of how different NLP-TM methods contribute to the analysis of sustainability reports.

In summary, the figures illustrate a diverse methodological landscape in SR analysis. The reliance on medium to advanced methods underscores their ability to balance complexity with analytical capability, while the preference for surface-level techniques in certain objectives highlights a focus on practicality and ease of use. This comparative analysis provides valuable insights into the

strengths and gaps of various methodologies, guiding future research in selecting the most appropriate techniques for specific analytical needs in SR analysis.

Figure 5: Mapping of research objectives, NLP methods, and analysis tools used to evaluate SRs (intersection view)



In the following discussion, we elaborate on each research objective and explore how specific methods and techniques are utilized to meet these goals, thereby enriching our understanding of the dynamic field of sustainability reporting analysis.

4.3.1 Identify themes and patterns in SRs

Studies in this category employed a variety of text-mining techniques to extract topics and themes from SRs, including supervised and unsupervised topic modelling methods, keyword extraction, term frequencies, and dictionary-based approaches. Table 2 provides detailed information on each study, including the extracted themes, methods used, and data scope. The primary objective of these studies was to uncover hidden patterns in how companies disclose sustainability information across different industries and sectors.

Overview of Key Methods

LDA was extensively used to reveal recurring themes in SRs, providing deep insights into corporate sustainability disclosures. This method effectively identified latent topics within the text, such as "environmental impact" and "corporate governance," demonstrating its robustness in thematic analysis (Benites-Lazaro et al., 2018; Hadro et al., 2022; Jaworska & Nanda, 2018; Mangsor et al., 2022; Ning et al., 2021; Niveditha et al., 2020; Rhoden et al., 2023; Szekely & Vom Brocke, 2017; Tiscini et al., 2021; Yang & Yang, 2022; Y. Zhou et al., 2022). LSA transformed unstructured text into structured data, enabling the extraction of themes by creating term-by-document matrices. This method was effective in analyzing term relationships and latent structures, although it may miss some contextual nuances compared to LDA. It was effective for meaningful topic extraction and trend analysis, revealing how terms like "renewable energy," "carbon footprint," and "sustainability initiatives" frequently appeared together (Kountouri et al., 2019; Reuter et al., 2014; Ye et al., 2022).

BERT, as a transformer-based technique, captured contextual information and generated accurate representations of text. BERT excelled in understanding context and nuances within SRs, making it highly effective for detailed analysis and entity recognition (Devlin et al., 2018; Ehrhardt & Nguyen, 2021; Grootendorst, 2022). However, its application was limited to smaller datasets due to high computational demands.

Keyword extraction and term frequency techniques identified relevant terms in SRs (Calabrese et al., 2023; Klimczak et al., 2023; Liew et al., 2014; Modapothala & Issac, 2009; Saha & Nabareseh, 2015; Uyar et al., 2021). Keyword extraction methods such as keyword matching and dictionary-based approaches pinpointed specific terms of interest. Modapothala & Issac (2009) Used a "bag-of-words" approach to categorize SRs based on sustainability stages. Term frequency analysis methods, like TF-IDF, measured the importance of terms in documents. Whittingham et al. (2022) applied TF-IDF to analyze the language of the United Nations Sustainable Development Goals (SDGs) and Ding et al. (2023) utilized it to estimate climate-related disclosures.

Clustering techniques identified similarities between documents or text segments, facilitating related theme grouping. For instance, Tremblay et al. (2015) used supervised machine learning with neural networks to identify sustainability patterns while Raghupathi et al. (2020) implementing the K-Nearest Neighbour (KNN) method for topic classification. Unlike topic modeling methods like LDA, clustering focuses on grouping existing data points based on similarity rather than uncovering latent topics.

Various software tools, such as Leximancer, NVivo, and CFIE, were employed for thematic analysis and pattern extraction, offering both qualitative and large-scale text analysis capabilities. For example, Albitar et al. (2021) used CFIE to analyze COVID-19 disclosures in SRs, while Corazza et al. (2020) employed Sketch Engine and Amazon's Rekognition for text and image analysis in corporate reports.

Table 2 provides a detailed overview of each SR study, including the methods used, study timeframe, dataset size, and specific techniques, offering a comprehensive view of how each approach was applied. Table 3 then groups studies by their use of the same method, comparing each method's inherent computational capabilities (Column 3) with the maximum dataset breadth covered across studies. This allows for an assessment of potential gaps between a method's theoretical strengths from a computer science perspective and its practical application in SR research.

Table 2: Theme and Pattern Extraction

Author(s), Year	Text Mining Category	Data Scope	Timespan
(Calabrese et al., 2023)	Topic modeling, regular expression, and frequency analysis	1,501	2016-2020
(Jafari et al., 2022)	Topic modeling, LDA, word frequency	120	2015-2022
(Raman et al., 2020)	Supervised Bert model training	125	NA
(Fiandrino & Tonelli, 2021)	Topic modeling, LDA	101	NA
(Raghupathi et al., 2020)	Topic modeling, machine learning-based text analytics and clustering based methods (KNN)	NA	2009-2019
(Sharma et al., 2022)	Topic modeling, extracting important bigrams	360	2019
(Benites-Lazaro et al., 2018)	Topic modeling LDA	NA	2016-2019
(Ye et al., 2019)	Bag of words, key word extraction	369	2011-2017
(Liew et al., 2014)	Topic modeling, keyword extraction, TFID, bigram/trigram	112	2011
(Saha & Nabareseh, 2015)	Topic modeling, the frequency of terms	8	2012-2013
(Uyar et al., 2021)	Topic modeling, key word extraction	478	1999–2018
(Zhou et al., 2021)	Topic modeling, LDA	33	2016-2019
(Ning et al., 2021)	Topic modeling, LDA	680	3 years
(Li & Zhao, 2021)	Topic modeling, LDA	181	2001-2018

(Szekely & Vom Brocke, 2017)	Topic modeling, LDA	9,514	1999-2015
(Tremblay et al., 2015)	Topic modeling, Singular Value Decomposition (SVD), and supervised learning using a neural network	17	2004-2012
(Pan, 2016)	Topic modeling, LSA	NA	NA
(Reuter et al., 2014)	Topic modeling, LSA	4,999	1999-2013
(Zakaria et al., 2021)	Dictionary-based	140	2015-2017
(Deng et al., 2017)	Dictionary-based	449	2001-2015
(Niveditha et al., 2020)	LDA, TF-IDF, term frequency, random forest, decision tree, SVM	NA	NA
(Goloshchapova et al., 2019)	LDA and manual analysis	3,618 reports	1999 to 2017
(Corazza et al., 2020)	lexicological analysis and machine learning image classification.	NA	2010 -2016
(Ignatov, 2021)	Term frequency, TF-IDF	11,000	2013-2018
(Rhoden et al., 2023)	LDA	113 reports	2013 to 2020
(Whittingham et al., 2022)	TF-IDF, Term frequency	440 reports	2011-2019
(Ding et al., 2023)	TF-IDF	2,659	2017-2019
(Kvasničková Stanislavská et al., 2023)	STM	2,100	2020
(Hummel et al., 2022)	Cosine similarity	5,939	2008-2019
(Hadro et al., 2022)	LDA	46	2019
(Chen & Bouvain, 2009)	Leximancer software	34	NA
(Tate et al., 2009)	Crawdad software	100	NA
(Baier et al., 2020)	Term frequency and a term-document matrix	100	2016-2020
(Tiscini et al., 2021)	LDA	26	2014-2019
(Liu et al., 2022)	Word2vec	4,718	2016-2020

(Ho et al., 2022)	TF-IDF	NA	2012-2016
(Breijer & Orij, 2022)	Computer-Aided Text Analysis (CATA)	around 775	2012-2020
(Mangsor et al., 2022)	LDA	75	2017-2019
(Klimczak et al., 2023)	Term frequency	175	2017-2021
(Pollach, 2018)	CATA	744	2001-2010
(Kountouri et al., 2019)	LSA	188	2001-2013
(Sambhanthan & Potdar, 2016)	NVivo software	48	2015
(Sambhanthan & Potdar, 2015)	NVivo software	NA	2012-2014
(Albitar et al., 2021)	NVivo, CFIE software	NA	2020
(Jaworska & Nanda, 2018)	LDA	317	2000-2013
(Samkin, 2012)	CATPAC II software, artificial neural network	5 reports	2007-2011
(Yang & Yang, 2022)	LDA	97	2011-2020
(Shahi et al., 2012)	Naive Bayes classification algorithm	More than 1,000	NA
(Shahi et al., 2012)	Naive Bayes classification algorithm	More than 1,000	NA
(Al-Shaer & Hussainey, 2022)	Bag of words, term frequency	280	2014-2018
(Ehrhardt & Nguyen, 2021)	BERT, NER	31	NA
(Zhou, 2021)	STM, semantic network analysis	406	2017-2019

Table 3: Theme and pattern extraction methods

Method/Technique	Studies	Inherent Capabilities	Utilized Capabilities in SR analysis
LDA	Benites-Lazaro et al. (2018), Hadro et al. (2022), Jaworska & Nanda (2018),	**Depth:** High **Breath capacity**: Large datasets (up to millions)	**Breadth:** LDA was generally applied to medium-sized datasets (70-200 SRs), with a few studies extending its use to larger datasets (up to 9,514 SRs). This indicates possibilities for broader application.

	Mangsor et al. (2022), Ning et al. (2021), Niveditha et al. (2020), Rhoden et al. (2023), Szekely & Vom Brocke (2017), Tiscini et al. (2021), Yang & Yang (2022), Zhou et al. (2022), Jafari et al. (2022), Fiandrino et al. (2020), Li & Zhao (2021), Goloshchapova et al. (2019)		**Depth:** LDA is employed to uncover latent topics within SRs, revealing patterns such as "environmental impact" and "corporate governance." It provides substantial depth in understanding thematic structures.
LSA	Kountouri et al. (2019), Pan (2016), Reuter et al. (2014)	**Depth:** Medium to advance **Breath capacity:** Medium (Hundreds to thousands)	**Breadth:** LSA was applied to a minimum of 17, with most studies analyzing between 100 to 1000s SRs**,** with one study analyzing up to 4,999 SRs. **Depth:** LSA provided a detailed analysis of term relationships and latent structures in SRs. It was effective for transforming unstructured text into structured data, allowed for meaningful topic extraction and trend analysis. For instance, revealed that terms like "renewable energy," "carbon footprint," and "sustainability initiatives" frequently appear together in similar contexts, suggested they are thematically linked. Compared to LDA, LSA focuses more on the co-occurrence and semantic relationships between terms, providing a different angle on thematic analysis.
BERT	Ehrhardt & Nguyen (2021), Raman et al. (2020)	**Complexity:** High **Data handling capacity:** Medium (Hundreds to thousands)	**Breadth:** Applied to relatively small datasets (31 and 125 SRs), suggesting that its high computational demands may limit scalability. **Depth:** BERT exceled in understanding context and nuances within SRs, made it highly effective for detailed analysis and entity recognition. Studies used BERT achieved a high level of accuracy and depth in theme extraction.
Cosine Similarity	Hummel et al. (2022)	**Complexity:** Medium to advance **Data handling capacity:** Medium to large datasets (thousands to millions)	**Breadth:** Applied to large dataset of 5,939 SRs. **Depth:** Cosine similarity effectively measured the extent to which SRs discussed predefined topics by comparing text similarity. It provided valuable insights into the alignment of reports with specific themes, though it may lack the depth of more sophisticated models like LDA or BERT.
Word2Vec	Liu et al. (2022)	**Complexity:** High **Data handling capacity:** Large datasets (up to millions)	**Breadth:** Applied to large datasets of 4,718 SRs. **Depth:** Word2Vec captured semantic relationships between words, offered a detailed understanding of themes in SRs. It was effective for identifying primary focus area. Studies used Word2Vec achieved deep semantic insights but may face limitations in processing smaller datasets. (Word2Vec provides deep semantic insights but requires substantial data for accurate modeling, making it effective for large-scale analyses.)
Software-Based Techniques	Breijer & Orij (2022), Pollach (2018), Mani et al. (2018), Sambhanthan & Potdar (2015, 2016), Samkin (2012), Kvasničková Stanislavská et al. (2023), Zhou (2021), Chen & Bouvain (2009), Tate et al. (2009), Albitar et al. (2021)	**Complexity:** Surface Level to Medium **Data handling capacity:** medium to large datasets (hundreds to thousands to millions)	**Breadth:** The number SRs analyzed varied widely. Some studies analyzed a small number of reports (e.g., 8, 26, 34), while others analyzed hundreds to thousands (e.g., up to 2,100 reports). **Depth:** These software-based techniques offered a range of capabilities from large-scale text analysis and pattern identification to detailed qualitative insights and dynamic topic modeling.
Supervised techniques;	Shahi et al. (2012), Corazza et al. (2020) , Tremblay et al. (2015)	**Complexity:** Low to high **Data handling capacity:** medium to Large (Thousands to millions)	**Breadth:** Studies using supervised techniques analyzed SRs rom relatively small (e.g. 10 SRS to around 100 SRs) to thousands of reports with one study analyzing over 1,000 reports using Naive Bayes classification. **Depth:** classification methods categorized text based on predefined labels, offered a straightforward approach to pattern recognition and theme extraction. They were effective for classifying textual data into specific categories based on content.
Word Frequency Analysis: Dictionary based; TF-IDF	Klimczak et al. (2023), Calabrese et al. (2023), Sharma et al. (2022), Ye et al. (2019), Saha	**Complexity:** Low **Data handling capacity:** Medium (Hundreds to thousands)	**Breadth:** The number of SR analyzed was relatively high compared to other methods, ranged from hundreds to over 11,000, with most studies focusing on medium-sized datasets (hundreds to thousands of reports). **Depth:** Term frequency and dictionary-based analysis provided

	& Nabareseh, (2015), Uyar et al. (2021), Baier et al. (2020), Al-Shaer & Hussainey (2022), Whittingham et al. (2022), Ding et al. (2023), Ho et al. (2022), Liew et al. (2014), Ignatov, 2021, Zakaria et al. (2021), Deng et al. (2017)		preliminary insights into frequently discussed topics in SRs. It highlighted the most common terms Studies used this technique were generally surface-level; TF-IDF was useful for identifying important terms and their relevance in SRs. It provided a quantifiable basis for term significance but did not capture deeper contextual relationships. Studies used TF-IDF often achieved a surface-level understanding of themes.
KNN clustering	Raghupathi et al. (2020)	**Complexity:** Medium to advance **Capacity:** medium to large	**Breadth:** The number of SR analyzed was not clear. **Depth:** KNN employed for pattern recognition and clustering, which can provide a medium level of depth in identifying thematic patterns and similarities within the data.

Table 3 illustrates that LDA, the most frequently used method for this objective, was applied to a range of dataset sizes. While its high depth allows for deep insights into latent topics within the text, it was generally applied to medium-sized datasets (70 to 200 SRs). Notably, only a few studies utilized LDA for large datasets, indicating a promising area for broader application. The study analyzing 1500 and 9,514 SRs demonstrated LDA's capability to handle substantial data volumes, revealing its capacity for scalability (Calabrese et al., 2023; Szekely & Brocke ,2017).

LSA was used in studies analyzing medium to large datasets, with one study handling up to 4,999 SRs (Reuter et al., 2014). LSA provides a detailed analysis of term relationships and latent structures in SRs. It was effective for transforming unstructured text into structured data, allowing for meaningful topic extraction and trend analysis. However, it may miss some contextual nuances compared to LDA.

BERT, though highly effective for detailed analysis and entity recognition, was applied to relatively smaller datasets (31 and 125 SRs). Its high computational demands might limit its scalability, suggesting that while BERT offers substantial depth, its application breadth is narrower compared to other methods like LDA.

Surface-level methods like Cosine Similarity and Word Frequency Analysis were applied to larger datasets. Cosine Similarity was used on a dataset of 5,939 SRs, effectively measuring text similarity and providing insights into the alignment of reports with specific themes. Similarly, TF-IDF was used to analyze term significance across documents, handling up to 2,659 reports. These methods, while providing broader applicability due to lower computational demands, offer less detailed insights compared to advanced techniques like LDA and BERT.

Software-based techniques like Nvivo and CATA were used in studies with varying dataset sizes, from small (8 reports) to large (up to 2,100 reports). In comparison with other methods in terms of

deph breath and scalability witnessed in studies, these methods balance depth and breadth, enabling both large-scale text analysis and detailed qualitative insights. Nvivo captured complex themes and patterns, offering comprehensive qualitative insights, while CATA enabled large-scale text analysis, identifying recurring themes and trends over time.

Overall, LDA and LSA emerged as the most frequently used methods, offering substantial depth and breadth in text analysis. While these methods were predominantly applied to small to medium-sized datasets, there are instances where they were used to their full computational capacity, handling large datasets effectively. BERT, despite its capability for detailed analysis, was underutilized due to its high computational demands, and was generally applied to smaller datasets. Simpler methods like TF-IDF, Cosine Similarity, and Term Frequency Analysis offered broader but shallower insights, making them more suitable for preliminary analyses and large datasets. This synthesis underscores the strengths and limitations of each method, providing valuable guidance for researchers in selecting the appropriate technique based on their specific analysis needs.

4.3.2 Assess compliance with SR frameworks

The second group (see details of studies in Table 4) employed TM-NLP methods to investigate the completeness of the information published in the SRs based on pre-defined non-financial reporting standards and frameworks (e.g., GRI, SASB, SDGs)[1] (Aureli, 2017; Aureli et al., 2016; S.-H. Liu et al., 2017; Shahi et al., 2014; Tóth et al., 2022; X. Wang et al., 2020). These studies collectively apply different techniques to evaluate SRs, assess their alignment with GRI, SASB, and SDGs, analyze ESG scores, evaluate sustainability strategies, and assess green performance in various sectors. These methodologies contribute to a deeper understanding of sustainability practices, disclosure standards, and the relationship between financial and non-financial information in corporate reporting.

[1] GRI was founded in Boston, USA in 1997 with the involvement of the United Nations Environment Programme (UNEP). It was a spin-off from the US non-profit organizations the Coalition for Environmentally Responsible Economies (CERES) and the Tellus Institute. GRI is one of the most widely used reporting format for CSR or sustainability reports worldwide (Ernst & Young, 2016). Initially GRI reporting frameworks were developed with investors at its core but over the course of time it took a multi-stakeholder stance. The frameworks thus broadened to include social, economic, and governance issues. SASB, on the other hand, is a rather contemporary institution founded in 2011. The SASB has currently issued 77 industry specific standards as of 2018, November. It is an independent 501(c)3 institution accredited by the American National Standards Institute (ANSI) and established to develop and disseminate sustainability accounting standards. The aim of the SASB is to provide standards which has decision usefulness, captures material aspects, and are cost effective. The standards are built upon evidence, market consensus and are industry specific. SASB takes a investors approach therefore standards are expected to carry information of material significance to shareholders.

Several studies utilized supervised ML techniques and text-mining software to analyze SRs for compliance with the GRI framework. Shahi et al. (2014) established a training corpus and categorized SRs into performance indicators, testing various document classification algorithms to determine the most accurate methods for categorizing text sections. Similarly, Chang & Cheng (2013) applied text mining to segment Chinese words in SRs, counting their frequencies to reveal differences and similarities in corporate sustainability emphases across sectors.

Unsupervised ML approaches were also employed to assess compliance with GRI standards. Gutierrez-Bustamante & Espinosa-Leal (2022) developed a matching index algorithm using exploratory data analysis, information retrieval, and NLP techniques, while Aureli (2017) and Aureli et al.(2017) used text mining to analyze the impact of industrial disasters on SR disclosures, employing glossaries based on GRI guidelines to quantify changes in disclosure.

Studies also focused on the alignment of SRs with SASB and SDGs. Lindgren et al. (2021a) proposed a novel approach to test SR alignment with SASB standards by converting the SASB materiality matrix into categories and performing LDA topic modeling. Tóth et al.(2022) explored the connection between SR and financial reporting, discussing the role of sustainability in financial reporting and highlighting efforts to standardize sustainability reporting. Tóth et al. (2021) analyzed disclosure requirements by automakers, employing keyword frequencies and TF-IDF to evaluate compliance with GRI and SASB standards.

Various methodologies were used to analyze the alignment between SRs and SDGs. Wang et al. (2020) manually assigned paragraphs from SRs to specific SDGs and employed association-based text mining to analyze the relationships between sustainability topics and SDGs. Angin et al. (2022) combined ML and deep learning models to automate the alignment of text blocks with SDGs, using text pre-processing, vectorization, and models like BERT and RoBERTa. Caliskan et al. (2022) used NVivo to analyze responsible consumption and production dimensions aligned with SDG 12, ranking companies based on their performance in these areas.

Other studies evaluated the effectiveness of SRs in communicating sustainability strategies and their impact on corporate operations.Vaio et al. (2022) assessed the alignment of SRs with non-financial reporting standards and frameworks, using Leximancer for automated text analysis and manual content analysis for waste management practices. Aguado-Correa et al. (2023) analyzed non-financial information disclosure in the Spanish banking sector using manual content analysis, NLP-based text scanning, and the TOPSIS methodology to rank entities based on their SDG contributions.

In exploring the relationship between financial statements and ESG scores, D'Amato et al.(2021) employed regression tree-based methods and random forest techniques to analyze the relationship between predictor variables and ESG scores. Bodendorf et al. (2022) used the word2vec model to quantify SWOT categories in SRs and evaluate sustainability strategies in the automotive industry. Ramakrishnan et al.(2023) developed a green performance assessment framework for airport buildings, using decision tree-based models to relate green ratings with airport features.

Finally, Moreno & Caminero (2022) analyzed TCFD recommendations on climate-related disclosures by Spanish financial institutions, using Python scripts for text extraction and manual review to create a taxonomy linked to a lexicon, facilitating a comprehensive analysis of climate-related disclosures.

These studies highlight the diverse application of TM-NLP methods in evaluating the completeness of SRs, providing valuable insights into how companies align their disclosures with established standards and the broader implications for corporate sustainability practices.

Table 4: Analysis of compliance with existing frameworks

Author(s), Year	**Text Mining Category**	**Data Scope**	**Timespan**
(Wang et al., 2020)	software RapidMiner, association-based text mining	56	2016-2019
(Aguado-Correa et al., 2023)	An open-access software (unsupervised text mining techniques, 2030 Agenda)	12	2020
(Moreno & Caminero, 2022)	Key word extraction and key word searching	118	2014-2020
(Gutierrez-Bustamante & Espinosa-Leal, 2022)	Unsupervised learning (LSA/LDA, Word2Vec)	524	2020
(Vaio et al., 2022)	Leximancer software which use Bayesian theory and unsupervised methods	18	2015-2019
(Tóth & Suta, 2021)	keyword frequencies and TF*IDF	158	2016-2020
(Tóth et al., 2022)	word and phrase frequency calculation, topic formation with factor analysis	NA	2019
(Ramakrishnan et al., 2023)	Cart model	577	2000-2021
(Caliskan et al., 2022)	NVivo 11 Plus software	16	2018-2019

(Bodendorf et al., 2022)	word2vec model	66	2019
(Modapothala & Issac, 2009)	Keyword frequency	2,415	2008
(Angin et al., 2022)	Supervised learning (RoBERTa)	NA	NA
(D'Amato et al., 2021)	Regression tree, random forest,	109	2014-2018
(Tóth et al., 2021)	Provalis WordStat software	NA	2016-2020
(Lindgren et al., 2021)	Topic modeling, LDA	9,500	1998–2017
(Kiriu & Nozaki, 2020)	Combination of Supervised learning (Neural network) and Unsupervised learning (Hierarchical Clustering)	8,729	1999-2016
(Liu et al., 2017)	Unsupervised model (DBSCAN), Clustering and Network analysis	50	NA
(Duan et al., 2018)	Part of Speech (POS)	2,970	2009-2015
(Aureli et al., 2017)	Dictionary-based, Text Clustering	10	2009-2013
(Shahi et al., 2014)	Supervised learning, including Naïve Bayes, Decision Table, Random Subspace, and Neural Networks.	593	2010
(Chang & Cheng, 2013)	Term frequency	NA	NA
(Aureli et al., 2016)	Dictionary-based, word frequency	20	5 years
(Ye et al., 2022)	LSA topic modeling	68	2011-2017

Table 5 summarizes the various methods and techniques used in SR studies within this group, detailing their computational complexity (depth) and data capacity (breadth) from a computer science perspective. It also evaluates how these aspects were utilized and implemented to achieve the research objectives.

Table 5: Methods to Assess compliance with standards

Method/Technique	Studies	Inherent Capabilities	Utilized capabilities in SR analysis
Supervised Techniques;	Shahi et al. (2014), Angin et al. (2022), D'Amato et al. (2021), Kiriu & Nozaki (2020), Ramakrishnan et al. (2023)	**Complexity:** High **Data Handling Capacity:** Medium to large datasets	**Breadth:** Supervised ML techniques were applied to datasets ranging from 109 to 8,729 SRs. Most studies analyzed medium datasets, with important examples being 593 SRs and 8,729 SRs. **Depth:** Provides detailed compliance assessment with high accuracy, capturing nuanced details. Effective for large-scale analyses but requires extensive computational resources.
Unsupervised Techniques; LDA, LSA, Clustering; word2vec, Association-Based Text Mining	Gutierrez-Bustamante & Espinosa-Leal (2022), Liu et al. (2017), Lindgren et al. (2021), Bodendorf et al. (2022), Ye et al. (2022), Wang et al. (2020)	**Complexity:** medium to High **Data Handling Capacity:** Medium to large datasets	**Breadth:** Unsupervised ML techniques were applied to a minimum of 50 SRs, with most studies analyzing between 50 to 70 SRs. Only one study analyzed a large dataset of 9,500 reports. **Depth:** Provided detailed compliance assessment with high accuracy and captured nuanced details. Were effective for large-scale analyses; Association-based text mining was useful for identifying associations between terms related to compliance. Provided preliminary insights into term associations.
Software-Based Techniques	Vaio et al. (2022), Caliskan et al. (2022), Aguado-Correa et al. (2023), Tóth et al. (2021)	**Complexity**: Low to Moderate **Data Handling Capacity:** Small to medium datasets (up to hundreds)	**Breadth:** Software-based techniques were applied to small datasets ranging from 12 to 18 SRs. **Depth:** Offered comprehensive qualitative analysis of compliance themes.
Keyword Frequency, Dictionary-Based, and TF-IDF, PoS	Chang & Cheng (2013), Aureli et al. (2016), Aureli (2017), Tóth et al. (2021), Tóth et al. (2022), Moreno & Caminero (2022), Duan et al. (2018), Modapothala & Issac (2009),	**Complexity:** Low (Surface Level**)** **Data Handling Capacity:** Medium to large datasets	**Breadth:** These techniques were applied to datasets ranging from 10 to 2,970 SRs. Most studies analyzed medium to large datasets, with important examples being 2,415 SRs. **Depth:** Provided quantitative insights and preliminary thematic analysis. Were effective for assessing compliance with predefined standards.

The table illustrates that advanced techniques such as supervised and unsupervised ML methods offer substantial depth and breadth in analyzing SRs for compliance with various standards. These methods were applied to medium to large datasets, providing detailed compliance assessments with high accuracy.

Supervised ML techniques, such as CART, BERT, regression trees, neural networks, and Naïve Bayes, were applied to medium to large datasets, ranging from 109 to 8,729 SRs. These methods provide detailed compliance assessments with high accuracy, capturing nuanced details. However, they require extensive computational resources and time for training.

Unsupervised ML techniques, such as those employed by Gutierrez-Bustamante & Espinosa-Leal (2022) and Liu et al. (2017), do not require labeled training data and are particularly useful for exploring data to identify underlying structures and themes. These methods were typically applied to smaller datasets, with most studies analyzing between 50 to 70 SRs, offering broad insights into compliance without the need for predefined categories. For instance, Liu et al. (2017) used LDA to model topics in SRs and assess their alignment with SASB standards, revealing complex compliance patterns and providing nuanced insights. Association-based text mining was particularly useful for identifying associations between terms related to compliance.

Software-based techniques like Leximancer and NVivo, as used by Vaio et al. (2022) and Caliskan et al. (2022), demonstrated low to moderate complexity and were applied to small to medium datasets. These methods offer comprehensive qualitative analysis of compliance themes but were limited by the size of the datasets they were applied to, ranging from 12 to 18 SRs. These methods are particularly effective for initial compliance assessments, as they streamline the process of thematic analysis, making it easier to identify key themes and trends in SRs without the intensive computational requirements of advanced ML techniques. However, they were applied to smaller datasets, limiting their overall breadth.

Simpler methods such as term frequency, TF-IDF, and dictionary-based approaches, utilized by studies like Chang & Cheng (2013) and Aureli et al. (2016), provide surface-level insights suitable for preliminary analyses and large datasets. These methods are effective for assessing compliance with predefined standards, offering quantitative insights but possibly missing nuanced details. For example, Aureli et al. (2017) used dictionary-based methods to quantify changes in disclosure, providing useful quantitative insights but lacking deeper contextual understanding.

Overall, while all methods ultimately provide insights regarding compliance with different standards from SRs, the key difference lies in the depth of the insights they offer and the breadth they can manage based on computational resources. Advanced techniques like supervised and unsupervised ML (e.g., Neural networks, BERT, LDA, Word2Vec) offer a more in-depth and nuanced understanding of the text, capturing semantic relationships and providing detailed contextual analysis. These methods are particularly effective for uncovering complex patterns and compliance nuances, as seen in studies by Shahi et al. (2014) and Lindgren et al. (2021), which analyzed large datasets and provided detailed compliance assessments.

In contrast, simpler methods like term frequency, TF-IDF, and dictionary-based approaches offer more straightforward, surface-level insights, making them suitable for preliminary analyses and large datasets but less effective for capturing deeper, contextual meanings. The synthesis highlights the strengths and limitations of each method, guiding researchers in selecting the appropriate technique based on their analysis needs. These simpler methods are advantageous for handling large volumes of data quickly and providing preliminary insights, but they fall short of capturing the comprehensive, detailed themes that advanced techniques can reveal. The choice of method depends on the dataset size and specific depth required for compliance assessment, guiding researchers in selecting the appropriate technique based on their analysis needs.

4.3.2 Analyze the readability and sentiment of SR disclosures

The third group (detailed in Table 6) investigates how emotions and readability are expressed in SRs as determinants of the company's performance and challenges, providing insights into the linguistic tone and complexity of SRs (Albitar et al., 2022; Cho, 2009; Cho et al., 2010; P. Clarkson et al., 2008; Du & Yu, 2021; Feng & Gao, 2020; Harjoto et al., 2020; Harymawan et al., 2020; H. Kang & Kim, 2022; Liang & Wu, 2022; Rocca et al., 2020; Sai et al., 2019; Y. Zhou et al., 2022).

Several studies utilized sentiment analysis to examine the emotional tone of SRs. Harymawan et al.(2020) found that sustainability challenges are often represented using negative words, while positive words reflect companies' actions toward addressing these challenges. Sai et al. (2019) conducted sentiment analysis using unigram and bi-gram word clouds to analyze emotions in yearly reports. Kang & Kim (2022) employed the 'DistilBERT base uncased fine-tuned SST-2' model to calculate the positive–negative sentiment ratio of the content. This model, designed to be faster and lighter while retaining most of BERT's language understanding capabilities, assigns sentiment scores to sentences, categorizing them as positive or negative based on a threshold score. This approach allowed researchers to visualize sentiment patterns and tendencies in SRs over time.

Other studies focused on the use of sentiment analysis in different contexts. Rocca et al.(2020) analyzed social media content from Italian local government organizations' public Facebook pages, employing lexicon dictionaries and convolutional neural networks to compute the sentiment of citizens' comments. This study explored stakeholders' judgments and sentiment regarding CSR communications and actions. Similarly, Du & Yu (2021) quantified the proportion of positive and negative words in CSR reports to measure the tone and analyze its influence on future CSR performance and market reactions.

ML models were also utilized for sentiment analysis. Clarkson et al. (2020) used XGBoost and Random Forest classifiers to analyze the sentiment and readability of CSR reports, extracting valuable textual features to predict a company's CSR performance. Albitar et al. (2022) employed automated text analysis using the CFIE-FRSE app from Lancaster University on CSR narratives, utilizing different regression models to test hypotheses about the relationship between corporate governance and CSR tone.

Several studies used DICTION, a computerized text analysis software, for content analysis. Cho et al. (2010) and Feng & Gao (2020) used DICTION to analyze the language tone in environmental disclosures, focusing on "optimism," "realism," and "certainty." Harjoto et al. (2020) used DICTION

6.0 to analyze CSR reports, constructing measures of psychometric properties to investigate how narrative characteristics might be associated with the gender of report signers and influence future social performance.

Lexical sentiment analysis was also applied in some studies. Liang & Wu (2022) created a lexicon for Chinese based on CSR reports, identifying positive and negative words to determine the tone of the CSR reports. They used the Jieba Word Segmentation Tool for text preprocessing and calculated the percentage of positive and negative words in each CSR report.

TM-NLP methods were applied to measure the textual complexity of SRs. Smeunix et al. (2016) used NLP techniques, including part-of-speech tagging, named entity recognition, and syntactic parsing, to analyze the readability of SRs. They quantified key linguistic aspects of the text using readability formulae. Zhou et al. (2023) employed word segmentation and transformation of unstructured text data into word vectors to analyze Corporate Environmental Disclosure (CED) information, studying the relationship between corporate environmental performance and readability.

In summary, these studies contribute to the understanding of how emotions and readability are conveyed in SRs and their implications for company performance and challenges. By employing diverse TM-NLP methods, researchers provided valuable insights into the linguistic tone and complexity of SRs, highlighting the importance of clear and balanced communication in corporate sustainability reporting.

Table 6: Readability and Sentiment

Author(s), Year	Text Mining Category	Data Scope	Timespan
(Kang & Kim, 2022)	BERT model	60	2011-2020
(Rocca et al., 2020)	Word frequency, LDA	N/A	Until 2018
(Harymawan et al., 2020)	Sentiment analysis	152	2010-2018
(Sai et al., 2019)	Sentiment analysis	36	2015-2018
(Smeuninx et al., 2016)	Part-of-Speech tagging, word/syllables frequency	470	NA
(Albitar et al., 2022)	Dictionary based	NA	2008-2017
(Cho et al., 2010)	Diction software	190	2002
(Clarkson et al., 2020)	Supervised Ml models, random forest and XGBoost	2056	2002-2016

(Du & Yu, 2021)	Sentiment analysis, word frequency	More than 500	2002-2014
(Liang & Wu, 2022)	Sentiment analysis	NA	2006-2016
(Feng & Gao, 2020)	Diction software	1,566	2007-2012
(Zhou et al., 2023)	Word frequency	2289	2008-2019
(Harjoto et al., 2020)	Diction software	2060	2006-2015

Table 7: Methods for analyzing readability and sentiment

Method/Technique	Studies	Inherent Capabilities	Utilized Capabilities in SR analysis
BERT	Kang & Kim (2022)	**Complexity:** High **Data handling capacity:** Medium (Hundreds to thousands)	**Breadth:** Applied to a small dataset of 60 SRs **Depth:** Provided highly detailed and context-aware analysis, which were suitable for identifying complex sentiment patterns and relationships within SRs.
Sentiment Analysis	Harymawan et al. (2020); Sai et al. (2019); Du & Yu (2021); Liang & Wu (2022)	**Complexity:** Moderate **Data handling capacity:** Medium (Hundreds to thousands)	**Breadth:** Sentiment analysis techniques were applied to datasets ranging from 36 to over 500 SRs. Most studies analyzed medium datasets, with two notable examples being 152 SRs and over 500 SRs. **Depth:** Sentiment analysis provided insights into the tone and emotional content of SRs and was useful for understanding public perception and sentiment trends.
Supervised techniques	Clarkson et al. (2020);	**Complexity:** High **Data handling capacity:** Medium (Hundreds to thousands)	**Breadth:** Supervised ML models were applied to large datasets, with one notable study analyzing 2,056 SRs. **Depth:** Supervised ML models, including random forest and XGBoost classifiers, offered highly accurate and reliable results for compliance assessment. These models exceled in identifying specific metrics and themes from SRs when trained with well-labeled data.
Dictionary-based, word frequency, POS	Rocca et al. (2020); Smeuninx et al. (2016); Zhou et al. (2023)	**Complexity:** Low (surface level) **Data handling capacity:** Medium to large (thousands to millions)	**Breadth:** These methods are applied to large datasets ranging from 470 to 2,289 SRs. **Depth:** Dictionary-based methods provided a quick and straightforward analysis of predefined terms and themes. They were effective for large-scale surface-level evaluations.; Part-of-speech tagging and other NLP techniques were used for preprocessing in NLP tasks, helping to structure and analyze text.
Software	Cho et al. (2010), Feng & Gao (2020), Harjoto et al. (2020); Albitar et al. (2022)	**Complexity:** Low to Moderate **Data handling capacity:** Medium to large (thousands to millions)	**Breadth:** These techniques were applied to datasets ranging from 190 to 2,060 SRs. Most studies analyzed medium to large datasets, with notable examples being 1,566 SRs and 2,060 SRS. **Depth:** Provided automated thematic analysis.

The table above illustrates the different levels of depth and breadth that various TM-NLP methods offer in analyzing the readability and sentiment of SR disclosures. Advanced methods such as the BERT model provide highly detailed and context-aware analysis, capturing nuanced semantic relationships and offering accurate sentiment scores. For instance, Kang & Kim (2022) employed BERT to analyze sentiment in SRs, achieving high accuracy and depth. However, due to its high computational demands, BERT was applied to smaller datasets (e.g., 60 reports).

Supervised ML models like Random Forest and XGBoost classifiers, used by Clarkson et al. (2020) and Albitar et al. (2022), also offer high accuracy and reliability in sentiment and readability analysis. These models can handle larger datasets (up to 2,056 reports) but require substantial computational resources and well-labeled training data. Their ability to provide detailed and precise sentiment analysis makes them valuable for large-scale evaluations.

Moderate complexity methods such as general sentiment analysis techniques (Harymawan et al., 2020; Sai et al., 2019; Du & Yu, 2021; Liang & Wu, 2022) can handle medium to large datasets (up to 500+ reports). These methods offer valuable insights into the tone and emotional content of SRs, helping to understand public perception and sentiment trends. They are less computationally intensive than advanced models like BERT and provide broad overviews of sentiment.

Simpler, surface-level techniques such as dictionary-based methods (Albitar et al., 2022; Cho et al., 2010; Feng & Gao, 2020; Harjoto et al., 2020) offer quick and straightforward analysis of predefined terms and themes. These methods can handle large datasets (up to 2,289 reports) and are effective for large-scale-level evaluations but may miss nuanced details. Similarly, software tools like DICTION (Cho et al., 2010; Feng & Gao, 2020; Harjoto et al., 2020) provide automated thematic analysis, balancing depth and efficiency for medium to large datasets.

Part of speech tagging and other NLP preprocessing techniques, used by Smeuninx et al. (2016) and Zhou et al. (2023), are essential for structuring and analyzing text, helping to measure readability and linguistic complexity. These methods are applied to medium datasets (e.g., 470 reports) and serve as critical steps in comprehensive NLP tasks.

Overall, advanced techniques like BERT and supervised ML models offer deeper and more detailed insights into the sentiment and readability of SRs, capturing complex patterns and relationships within the text. These methods are particularly effective for nuanced and context-aware analysis but require substantial computational resources and well-labeled data. In contrast, simpler methods like dictionary-based approaches and software tools offer broader, surface-level insights, making them suitable for large-scale preliminary analyses but less effective for capturing the detailed, contextual nuances that advanced methods can reveal. The synthesis highlights the strengths and limitations of each method, guiding researchers in selecting the appropriate technique based on their analysis needs and the depth and breadth required for their studies.

5 Discussions

This systematic literature review aimed to explore the application of TM-NLP methods in the analysis of sustainability reports (SRs). By examining 88 selected publications, the study sought to answer two key research questions: (1) What are the different text mining methods employed to analyze SRs? and (2) What are the objectives of researchers applying these methods? The findings reveal three primary research objectives in SR analysis: the identification of SR themes and patterns, the assessment of SR compliance with existing frameworks, and the analysis of readability and sentiment in SR disclosures. The discussion below synthesizes the findings assesses the current state of the literature and suggests possible directions for future research.

Key Findings and Implications

Overall, our review indicates that there has been notable progress in leveraging TM-NLP techniques for the automated analysis of qualitative, unstructured data found in SRs. These advancements have undoubtedly contributed important insights into sustainability reporting. However, addressing our research questions has also highlighted areas that need further exploration and improvement.

- *In response to RQ1*, our findings reveal that the realm of SR analysis is composed by a rich tapestry of methodologies that range from simple word frequency analysis to cutting-edge NLP techniques like BERT (Devlin et al., 2018). Automated text analysis tools included in our review sample - topic modeling, clustering, sentiment analysis, and semantic-based processes - appear to be becoming vital to the in-depth and timely exploration of SRs. The adoption of a particular technique or a combination thereof is often influenced by the scale of the dataset and the depth of analysis sought. While vast datasets often attract more streamlined, automated methods, complex and nuanced analyses are typically reserved for smaller, more focused datasets. We explore this last point in considerable detail below.
- *In response to RQ2*, the objectives of employing TM-NLP in analyzing SRs, as per RQ2, are multifaceted. They range from understanding immediate linguistic structures and word frequencies to unralveling deeper, more nuanced emotions, strategies, and intentions buried within these reports.

Three objectives and methodologies identified are listed below:

1. *Identification of SR Themes and Patterns:* Advanced methods like LDA and BERT were predominantly used to uncover latent topics within SRs, offering deep insights into corporate

sustainability disclosures. These techniques facilitated the extraction of complex themes and patterns from large datasets. However, the computational demands of such methods often limit their application to smaller datasets, underscoring a important gap in scalability. Simpler methods like TF-IDF, which provide surface-level insights, were commonly used for initial thematic exploration and supported deeper analyses conducted by more sophisticated models

2. *Analysis of SR Compliance with Existing Frameworks:* The objective here was to evaluate the extent to which sustainability reports align with established non-financial reporting standards such as GRI, SASB, and SDGs. The review found that surface-level and software-based techniques were frequently employed, offering preliminary compliance assessments. Advanced methods were less commonly used, suggesting an opportunity for methodological enhancement. This indicates that while current methods are effective for preliminary assessments, they may not fully capture the nuanced compliance patterns that more sophisticated techniques could uncover.
3. *Sentiment and Readability Assessment:* Sentiment analysis and readability measures were applied to classify the sentiment expressed in SRs and measure language complexity. Techniques like BERT were notably effective in capturing nuanced sentiment and contextual relationships, though their application was often restricted to smaller datasets due to high computational demands. Other methods, including dictionary-based approaches and simpler sentiment analysis techniques, were employed to analyze larger datasets, providing broad overviews while being less resource-intensive.

The intersection of these findings raises some important issues that we suggest should be addressed going forward if the true capabilities of automated text analysis of SR are to be reached. We expand on these in the remainder of this section.

5.1 Tensions between depth and breath

A recurring theme we discovered in the literature is the trade-off between breadth and depth in SR analysis. Studies dealing with vast datasets have tended to prioritize breadth, using simpler methods like word frequency analysis or TF-IDF, which offer quick insights but may miss the deeper, more nuanced themes within the reports. Conversely, in-depth analyses that employed complex methodologies such as BERT or LDA tended to focus on smaller datasets, offering richer insights but limiting the generalizability of the findings.

This is important because simple techniques like term frequency and TF-IDF provide surface-level insights suitable for preliminary analyses and large datasets. However, they are limited when it

comes to analyzing the nuanced meanings in sustainability reports (SRs). For example, studies in the "Readability and Sentiment" section often relied on term frequency assessments, overlooking the complexities of SR content, and thus limiting the insights that can be garnered from the analysis.

In summary, while simple techniques are easier and quicker to operationalize and may prove valuable for preliminary analyses and large datasets, the advanced methods that are now on offer and outlined at the beginning of this paper, can offer much deeper insights. Future research should carefully consider this trade-off between breadth and depth. By doing so, researchers can make more informed decisions about the most suitable analytical methods for their specific research objectives.

5.1.1 Limitations in Longitudinal Analysis

Another important observation is the lack of longitudinal analyses in the reviewed studies. Although TM-NLP methods are well-suited for analyzing large datasets over extended periods, most studies have focused on a single point in time, overlooking the temporal dynamics of sustainability practices. This emphasis on static assessments limits our understanding of how SR content and sustainability practices evolve, which is essential given the rapidly changing ESG landscape and the inherent learning process in reporting (Hahn & Kühnen, 2013). By predominantly assessing sustainability outcomes at a specific moment, these studies miss the opportunity to conduct trend analyses that could reveal how sustainability practices develop and change over time.

This inclination towards static assessments is evident in several studies with a limited longitudinal scope. For instance, a notable number of studies analyzed the overall practice themes based on a static outcome, extracting themes, topics, and frequent words from SRs at a specific time point – for example, Liew et al. (2014) and Li & Zhao (2021), who extracted frequent words related to social themes from a set of reports, such as health and safety and employee development. See also Kvasničková Stanislavská et al. (2023) who only analyze the year 2020, Hadro et al.(2022) explore only the year 2019, and Mani et al. (2018) explore only the year 2015. This pattern is prevalent across our three overarching objective groups where studies focused on identifying specific words that indicate a specific sustainability theme or a firm's compliance with predefined guidelines.

A handful of studies stand out for their efforts to uncover temporal dynamics and trends in sustainability reporting Uyar et al. (2021);Szekely & Vom Brocke (2017). Uyar et al. (2021) made an effort to uncover patterns and changes in the topical structure of SRs in the Hospitality and Tourism industry over time, utilizing network analysis to analyze how the thematic structure of SRs evolved across different periods. However, even here it should be noted that some have based their trend

analysis on alternative data sources like media articles, rather than direct analysis of SRs. Similarly, Deng et al. (2017) explored environmental sustainability topics by mapping media coverage over time. However, their analysis was also based on New York Times articles rather than SRs. While they extracted the most frequent recent words from the media to construct a dictionary for analyzing SRs, their study did not involve a direct time series analysis of SRs. Moreover, while some research has hinted at temporal trends, such as studies by Aureli et al.(2017) that juxtaposed sustainability disclosures pre and post specific events. Yet, these often operate within limited timeframes or examine a relatively constrained number of reports.

Overall, while some studies such as Aureli et al.(2017) has touched upon the trend analysis perspective by comparing sustainability disclosures before and after specific events, they often suffer from limitations in terms of the short timeframes considered and the relatively small number of reports analyzed. This emphasis on static assessments limits the understanding of how SR and practice evolve over time and neglects the temporal analysis of sustainability performance. Many of the papers in our study have acknowledged this limitation and have highlighted the need to extend the adopted techniques to larger datasets, diverse industrial sectors, and different countries (Sai et al., 2019, Reuter et al., 2014, Tremblay et al., 2015, Pan, 2016, Saha & Nabareseh, 2015, Li & Zhao, 2021).

In summary, sustainable activities and reporting frameworks change over time and analyzing their development can provide valuable insights into emerging trends and patterns. Future research should incorporate dynamic perspectives and consider trend analysis models to capture this evolution. By taking a longitudinal approach, researchers can better understand the progress, changes, and challenges associated with sustainable development.

5.1.2 Need for breadth and depth in future research

The prevalent theme emerging from our exploration of the literature to date is clear - while extensive data exists, only a fraction of studies have realized the possibilities for automated text analysis approaches to generate results with substantial breadth *and* depth. We presume that the predominant use of basic methods for substantial datasets could be the result of challenges related to computational efficiency and technical manageability – analyzing big data with more advanced techniques requires a fairly deep and diversified skillset, as well as considerable computational resources. Likely this is because as these tools are relatively recent and require time and resources to be able to learn and utilize.

These observations underscore an important path for evolution in TM-NLP applications for SRs. As we advance, it becomes imperative to strike a balance between depth and breadth, ensuring that wide-ranging studies do not merely skim the surface but explore nuanced data layers for more profound insights. Similarly, the latent possibilities in investigating the broad datasets that cover multiple industries and time periods is vast, particularly when it comes to the temporal dynamics of sustainability practices. Unearthing patterns of change, advancements, and regressions in SR over time and sphere of activity can offer invaluable insights into the trajectory of sustainable practices and the broader socio-economic and environmental contexts influencing them. As the field progresses, it becomes paramount to integrate dynamic and cross-industry perspectives, diving into longitudinal studies to truly grasp the evolving trends and changes in SR practices over time.

5.2 Deficits in methodological rigor and transparency

Another important finding from our review is the lack of transparency and rigor in the methodological approaches of many studies. While some studies provided detailed descriptions of their methods, many did not justify their methodological choices or discuss the limitations of their chosen techniques. This lack of transparency can hinder the reproducibility and validity of research findings.

Moreover, comparative analyses of different methods were rarely conducted, which limits our understanding of the relative strengths and weaknesses of various TM-NLP techniques. Future research should emphasize methodological transparency and rigor, providing clear justifications for chosen methods and acknowledging their limitations. Additionally, comparative studies that evaluate multiple methods could help in identifying the most effective approaches for specific research objectives.

5.2.1 Justifying method and acknowledging limitations

One of the cornerstones of good science is the clear justification of the research method selected by a research paper, the steps followed to operationalize the method, and the limitations of the conclusions drawn from the findings (Pedhazur & Schmelkin, 2013). This good research practice provides a level of transparency that increases validity, allows for replication studies, and the cumulative advancement of knowledge (Gerring, 2012). Surprisingly, and a bit disappointingly, most of the studies in our original raw sample failed to provide a clear description of their methodological approach, thereby resulting in their being excluded from this SLR (see section 3.2 above in which we describe how we selected our sample for this review).

However, there is significant room for improvement even with the papers that were included into this review. Although they provided a description of their methods, many of these provided limited or no discussion of the suitability of the chosen method vis-à-vis alternative methodologies or comparative approaches. In addition, many studies included in our sample do not acknowledge the limitations or possible biases associated with the selected methods or introduced by the chosen techniques, such as the reliance on predefined dictionaries, subjectivity in topic modeling, or bias in sentiment analysis algorithms. We highlight the need to address these limitations in order to enhance the robustness and credibility of research findings.

5.2.2 Broader sets of tools

Comparative analysis can provide insights into the strengths and weaknesses of different methods (in terms of the depth they can analyze), help researchers make informed decisions about which technique is most suitable for their specific research objectives, or be implemented as a mixed method research design (Gerring, 2012; Pedhazur & Schmelkin, 2013). While some studies employed multiple methods within their analysis (e.g., combining topic modeling with keyword extraction) these were the exception.

Consequently, it seems clear that based on our findings there is lots of opportunity for further exploration of integrating various techniques to enhance the comprehensiveness and accuracy of the findings. For example, combining BERT with sentiment analysis or readability assessment could provide a more holistic understanding of sustainability reporting. This underscores the significance of combining various methodologies: while automation can speed up data extraction and initial analysis, a nuanced understanding typically emerges from the combination of various methods, both manual and automated (Kang, 2021; Liu et al., 2017). Embracing this blended approach ensures that while we benefit from the efficiency of automated text analysis tools, we also capture the richness and depth that can come from more sophisticated techniques and human interpretation.

6 Conclusions

Automated text analysis methods, such as text mining (TM) and natural language processing (NLP), hold great promise for managing the increasing volume, diversity, and complexity of data found in corporate sustainability reports (SRs). Despite this promise, our understanding of how these approaches are utilized in research remains limited due to their relative novelty. To address this gap, we conducted a systematic literature review (SLR) to synthesize and evaluate existing literature utilizing TM-NLP for SR analysis.

Our review provides a comprehensive overview of methodologies and techniques, examines research objectives, and critically assesses the progress made in terms of depth and breadth. While there has been notable progress, we have identified key issues related to research depth, breadth, rigor, and transparency, indicating that the considerable capability of TM-NLP to extract meaningful insights from SR big data has not been fully realized.

This literature review contributes to the existing body of knowledge by showcasing the promise of TM-NLP in analyzing SRs, providing a comprehensive understanding of the objectives and methods employed in previous studies and issues that should be addressed by further research. Overall, these findings have practical implications for organizations, policymakers, and stakeholders and facilitate benchmarking and sharing of best practices in SR and management.

References

Adam, C. A. (2022). *EU v IFRS: Fundamentally different approaches to sustainability reporting*. https://drcaroladams.net/eu-v-ifrs-fundamentally-different-approaches-to-sustainability-reporting/

Aggarwal, C. C., & Zhai, C. (2012). A Survey of Text Classification Algorithms. In C. C. Aggarwal & C. Zhai (Eds.), *Mining Text Data* (pp. 163–222). Springer US. https://doi.org/10.1007/978-1-4614-3223-4_6

Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. *Proc. 20th Int. Conf. Very Large Data Bases, VLDB*, *1215*, 487–499. https://www.it.uu.se/edu/course/homepage/infoutv/ht08/vldb94_rj.pdf

Aguado-Correa, F., de la Vega-Jiménez, J. J., López-Jiménez, J. M., Padilla-Garrido, N., & Rabadán-Martín, I. (2023). Evaluation of non-financial information and its contribution to advancing the sustainable development goals within the Spanish banking sector. *European Research on Management and Business Economics*, *29*(1), 100211. https://doi.org/10.1016/J.IEDEEN.2022.100211

Albitar, K., Abdoush, T., & Hussainey, K. (2022). Do corporate governance mechanisms and ESG disclosure drive CSR narrative tones? *International Journal of Finance & Economics*. https://doi.org/10.1002/IJFE.2625

Alliance for Corporate Transparency. (2020). Alliance for Corporate Transparency. https://www.allianceforcorporatetransparency.org/

Al-Shaer, H., & Hussainey, K. (2022). Sustainability reporting beyond the business case and its impact on sustainability performance: UK evidence. *Journal of Environmental Management*, *311*, 114883. https://doi.org/10.1016/J.JENVMAN.2022.114883

Angin, M., Taşdemir, B., Yılmaz, C. A., Demiralp, G., Atay, M., Angin, P., & Dikmener, G. (2022). A RoBERTa Approach for Automated Processing of Sustainability Reports. *Sustainability*, *14*(23), 16139.

Antons, D., Grünwald, E., & Cichy, P. (2020). The application of text mining methods in innovation research: Current state, evolution patterns, and development priorities—Antons—2020—R&D Management—Wiley Online Library. *R&D Management*, *50*(3), 329–351.

Aras, G., & Crowther, D. (2008). Corporate Sustainability Reporting: A Study in Disingenuity? *Journal of Business Ethics*, *87*(1), 279–288. https://doi.org/10.1007/S10551-008-9806-0

Auke, E., & Simaens, A. (2019). Corporate responsibility in the fast fashion industry: How media pressure affected corporate disclosure following the collapse of Rana Plaza. *International Journal of Entrepreneurship and Innovation Management*, *23*(4), 356–382. https://doi.org/10.1504/IJEIM.2019.100800

Aureli, S. (2017). A comparison of content analysis usage and text mining in CSR corporate disclosure. *International Journal of Digital Accounting Research*, *17*, 1–32. https://doi.org/10.4192/1577-8517-V17_1

Aureli, S., Medei, R., Supino, E., & Travaglini, C. (2016). Sustainability disclosure after a crisis: A text mining approach. *International Journal of Social Ecology and Sustainable Development*, *7*(1), 35–49. https://doi.org/10.4018/IJSESD.2016010102

Aureli, S., Medei, R., Supino, E., & Travaglini, C. (2017). Sustainability disclosure and a legitimacy crisis: Insights from two major cruise companies. *European Journal of Tourism Research*, *17*, 149–163.

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

Barbier, L. M., Green, J. L., & Draper, D. S. (2022). The need for open access and natural language processing. *Proceedings of the National Academy of Sciences*, *119*(15), e2200752119. https://doi.org/10.1073/pnas.2200752119

Barcellos de Paula, L., & Gil-Lafuente, A. M. (2021). Algorithm applied in corporate sustainability. *Kybernetes*, *50*(7), 1982–2000.

Bartholomew, D. J., Knott, M., & Moustaki, I. (2011). *Latent variable models and factor analysis: A unified approach*. John Wiley & Sons.

Baumüller, J., & Sopp, K. (2021). Double materiality and the shift from non-financial to European sustainability reporting: Review, outlook and implications. *Journal of Applied Accounting Research*, *23*(1), 8–28.

Beasley, M. S., Clune, R., & Hermanson, D. R. (2005). Enterprise risk management: An empirical analysis of factors associated with the extent of implementation. *Journal of Accounting and Public Policy*, *24*(6), 521–531. https://doi.org/10.1016/j.jaccpubpol.2005.10.001

Benites-Lazaro, L. L., Giatti, L., & Giarolla, A. (2018). Sustainability and governance of sugarcane ethanol companies in Brazil: Topic modeling analysis of CSR reporting. *Journal of Cleaner Production*, *197*, 583–591. https://doi.org/10.1016/J.JCLEPRO.2018.06.212

Betti, G., Consolandi, C., & Eccles, R. G. (2018). The relationship between investor materiality and the sustainable development goals: A methodological framework. *Sustainability*, *10*(7), 2248.

Billio, M., Costola, M., Hristova, I., Latino, C., & Pelizzon, L. (2021). Inside the ESG ratings: (Dis)agreement and performance. *Corporate Social Responsibility and Environmental Management*, *28*(5), 1426–1445. https://doi.org/10.1002/CSR.2177

Blasco, J. L. (2017). *The KPMG Survey of Corporate Responsibility Reporting 2017*.

Blazkova, T., Pedersen, E. R. G., Andersen, K. R., & Rosati, F. (2023). Greenwashing debates on Twitter: Stakeholders and critical topics. *Journal of Cleaner Production*, *427*, 139260. https://doi.org/10.1016/j.jclepro.2023.139260

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003a). Latent Dirichlet Allocation. *Journal of Machine Learning Research*.

Bodendorf, F., Dimitrov, G., & Franke, J. (2022). Analyzing and evaluating supplier carbon footprints in supply networks. *Journal of Cleaner Production*, *372*, 133601. https://doi.org/10.1016/J.JCLEPRO.2022.133601

Boiral, O., & Heras-Saizarbitoria, I. (2020). Sustainability reporting assurance: Creating stakeholder accountability through hyperreality? *Journal of Cleaner Production*, *243*, 118596. https://doi.org/10.1016/J.JCLEPRO.2019.118596

Brammer, S., & Pavelin, S. (2004). Building a good reputation. *European Management Journal*, *22*(6), 704–713.

Breijer, R., & Orij, R. P. (2022). The Comparability of Non-Financial Information: An Exploration of the Impact of the Non-Financial Reporting Directive (NFRD, 2014/95/EU). *Accounting in Europe*, *19*(2), 332–361. https://doi.org/10.1080/17449480.2022.2065645

Busco, C., Consolandi, C., Eccles, R. G., & Sofra, E. (2020). A Preliminary Analysis of SASB Reporting: Disclosure Topics, Financial Relevance, and the Financial Intensity of ESG Materiality. *Journal of Applied Corporate Finance*, *32*(2), 117–125. https://doi.org/10.1111/jacf.12411

Bushee, B. J., Gow, I. D., & Taylor, D. J. (2018). Linguistic Complexity in Firm Disclosures: Obfuscation or Information? *Journal of Accounting Research*, *56*(1), 85–121. https://doi.org/10.1111/1475-679X.12179

Calabres, A., Costa, R., Ghiron, N., & Menichini, T. (2019). Materiality analysis in sustainability reporting: A tool for directing corporate sustainability towards emerging economic, environmental and social opportunities. *Technological and Economic Development of Economy*, *25*, 1–23. https://doi.org/10.3846/tede.2019.10550

Calabrese, A., Costa, R., Ghiron, N. L., & Menichini, T. (2017). Materiality analysis in sustainability reporting: A method for making it work in practice. *European Journal of Sustainable Development*, *6*(3), 439–439.

Calabrese, A., Costa, R., Levialdi, N., & Menichini, T. (2016). A fuzzy analytic hierarchy process method to support materiality assessment in sustainability reporting. *Journal of Cleaner Production*, *121*, 248–264.

Calabrese, A., Costa, R., & Rosati, F. (2015). A feedback-based model for CSR assessment and materiality analysis. *Accounting Forum*, *39*(4), 312–327.

Calabrese, A., Costa, R., Tiburzi, L., & Brem, A. (2023). Merging two revolutions: A human-artificial intelligence method to study how sustainability and Industry 4.0 are intertwined. *Technological Forecasting and Social Change*, *188*, 122265. https://doi.org/10.1016/J.TECHFORE.2022.122265

Caliskan, A., Ozturkoglu, O., & Ozturkoglu, Y. (2022). Ranking of Responsible Automotive Manufacturers According to Sustainability Reports Using PROMETHEE and VIKOR Methods. *Advanced Sustainable Systems*, *6*(6), 2100301. https://doi.org/10.1002/ADSU.202100301

Cerbone, D., & Maroun, W. (2020). Materiality in an integrated reporting setting: Insights using an institutional logics framework. *The British Accounting Review*, *52*(3), 100876. https://doi.org/10.1016/J.BAR.2019.100876

Chang, D.-S., & Cheng, Y.-W. (2013). Explore the effects of industrial context and leaders' viewpoints on corporate sustainability in Taiwan by text mining. *Service Systems and Service Management (ICSSSM)*, 670–673. https://doi.org/10.1109/ICSSSM.2013.6602594

Chen, S., & Bouvain, P. (2009). Is corporate responsibility converging? A comparison of corporate responsibility reporting in the USA, UK, Australia, and Germany. *Journal of Business Ethics*, *87*(SUPPL. 1), 299–317. https://doi.org/10.1007/S10551-008-9794-0/METRICS

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016: 785-794. *ACM, New York, NY*.

Chiarini Tremblay William, M., Parra, C. M., & Castellanos William, A. (2015). *Corporate social responsibility reports: Understanding topics via text mining*.

Cho, C. H. (2009). Legitimation Strategies Used in Response to Environmental Disaster: A French Case Study of Total SA's Erika and AZF Incidents. *European Accounting Review*, *18*(1), 33–62. https://doi.org/10.1080/09638180802579616

Cho, C. H., Roberts, R. W., & Patten, D. M. (2010). The language of US corporate environmental disclosure. *Accounting, Organizations and Society*, *35*(4), 431–443. https://doi.org/10.1016/J.AOS.2009.10.002

Clark, K. P., Plc, A., Kotsantonis, S., Rehnberg, C., Ward, B., Advisors, K., Serafeim, G., Indahl, R., Gunvor Jacobsen, H., Equity, S., Davis, T., Lescott, B., Business School, H., Lins, K. V., Bizoumi, T., Lazaridis, S., Stamou, N., Ramani, V., Pinney, C., Capital Management, R. (2019). Four Things No One Will Tell You About ESG Data. *Journal of Applied Corporate Finance*, *31*(2), 50–58. https://doi.org/10.1111/JACF.12346

Clarkson, P., Li, Y., & Richardson, G. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. *Accounting, Organizations and Society*, *33*(4–5), 303–327.

Clarkson, P. M., Ponn, J., Richardson, G. D., Rudzicz, F., Tsang, A., & Wang, J. (2020). A Textual Analysis of US Corporate Social Responsibility Reports. *Abacus*, *56*(1), 1009–1051. https://doi.org/10.1111/ABAC.12182

Collins-Thompson, K. (2014). Computational assessment of text readability: A survey of current and future research. *ITL - International Journal of Applied Linguistics*, *165*(2), 97–135. https://doi.org/10.1075/itl.165.2.01col

Corazza, L., Truant, E., Scagnelli, S. D., & Mio, C. (2020). Sustainability reporting after the Costa Concordia disaster: A multi-theory study on legitimacy, impression management and image restoration. *Accounting, Auditing and Accountability Journal*, *33*(8), 1909–1941. https://doi.org/10.1108/AAAJ-05-2018-3488

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*(1), 21–27.

Crane, A., LeBaron, G., Allain, J., & Behbahani, L. (2019). Governance gaps in eradicating forced labor: From global to domestic supply chains. *Regulation & Governance*, *13*(1), 86–106. https://doi.org/10.1111/rego.12162

Cui, J., Wang, Z., Ho, S.-B., & Cambria, E. (2023). Survey on sentiment analysis: Evolution of research methods and topics. *Artificial Intelligence Review*, *56*(8), 8469–8510. https://doi.org/10.1007/s10462-022-10386-z

D'Amato, V., D'Ecclesia, R., & Levantesi, S. (2021). Fundamental ratios as predictors of ESG scores: A machine learning approach. *Decisions in Economics and Finance*, *44*(2), 1087–1110. https://doi.org/10.1007/S10203-021-00364-5

Delmas, M. A., & Burbano, V. C. (2011). The Drivers of Greenwashing. *California Management Review*, *54*(1), 64–87. https://doi.org/10.1525/cmr.2011.54.1.64

Deloitte. (2023). *A practical approach to assess financial materiality*. https://www2.deloitte.com/nl/nl/pages/sustainability/articles/a-practical-approach-to-assess-financial-materiality.html

Deng, Q., Hine, M., Ji, S., & Sur, S. (2017a). *Building an Environmental Sustainability Dictionary for the IT Industry*. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2017.112

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805*, *1*, 4171–4186. https://doi.org/10.48550/arXiv.1810.04805

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review*, *86*(1), 59–100.

DiMaggio, P. (2015). Adapting computational text analysis to social science (and vice versa). *Big Data & Society*, *2*(2), 205395171560290-. https://doi.org/10.1177/2053951715602908

Ding, D., Liu, B., & Chang, M. (2023). Carbon Emissions and TCFD Aligned Climate-Related Information Disclosures. *Journal of Business Ethics*, *182*(4), 967–1001. https://doi.org/10.1007/S10551-022-05292-X/TABLES/16

Du, S., & Yu, K. (2021). Do Corporate Social Responsibility Reports Convey Value Relevant Information? Evidence from Report Readability and Tone. *Journal of Business Ethics*, *172*(2), 253–274. https://doi.org/10.1007/S10551-020-04496-3/TABLES/9

Duan, G., Hu, W., Xue, H., & Lu, X. (2018). Financial risk assessment model of block chain based on soft computing and neural network. *IPPTA: Quarterly Journal of Indian Pulp and Paper Technical Association*, *30*(4), 108–118.

Eccles, R. G., & Krzus, M. P. (2010). *One report: Integrated reporting for a sustainable strategy*. John Wiley & Sons.

Eccles, R. G., Krzus, M. P., Rogers, J., & Serafeim, G. (2012). The Need for Sector-Specific Materiality and Sustainability Reporting Standards. *Journal of Applied Corporate Finance*, *24*(2), 65–71. https://doi.org/10.1111/j.1745-6622.2012.00380.x

Eccles, R. G., Serafeim, G., & Krzus, M. P. (2011). Market Interest in Nonfinancial Information. *Journal of Applied Corporate Finance*, *23*(4), 113–127. https://doi.org/10.1111/J.1745-6622.2011.00357.X

Ehrhardt, A., & Nguyen, M. T. (2021). Automated ESG Report Analysis by Joint Entity and Relation Extraction. *Communications in Computer and Information Science*, *1525 CCIS*, 325–340. https://doi.org/10.1007/978-3-030-93733-1_23/FIGURES/11

ESG Seneca, S. (2021, April 22). Sustainability Reporting: GRI and SASB Complement Each Other. *Seneca ESG*. https://senecaesg.com/insights/sustainability-reporting-gri-and-sasb-complement-each-other/

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd*, *96*(34), 226–231.

Fabrizio, K. R., & Kim, E.-H. (2019). Reluctant disclosure and transparency: Evidence from environmental disclosures. *Organization Science*, *30*(6), 1207–1231.

Fasan, M., & Mio, C. (2017). Fostering Stakeholder Engagement: The Role of Materiality Disclosure in Integrated Reporting. *Business Strategy and the Environment*, *26*(3), 288–305. https://doi.org/10.1002/bse.1917

Feldman, R., & Sanger, J. (2007). *The text mining handbook: Advanced approaches in analyzing unstructured data*. Cambridge university press.

Feng, S., & Gao, L. S. (2020). The Verbal Tone in Mandatory Environmental Disclosures: Evidence from Changes in Disclosures Following SEC Guidance. *Social and Environmental Accountability Journal*, *40*(2), 116–139. https://doi.org/10.1080/0969160X.2020.1719172

Fiandrino, S., & Tonelli, A. (2021). A Text-Mining Analysis on the Review of the Non-Financial Reporting Directive: Bringing Value Creation for Stakeholders into Accounting. *Sustainability 2021, Vol. 13, Page 763*, *13*(2), 763. https://doi.org/10.3390/SU13020763

Fiandrino, S., Tonelli, A., & Devalle, A. (2022). Sustainability materiality research: A systematic literature review of methods, theories and academic themes. *Qualitative Research in Accounting & Management*, *19*(5), 665–695. https://doi.org/10.1108/QRAM-07-2021-0141

Gamsjäger, L., & Ray, D. (2024, June 21). GRI, SASB, CDP – Making sense of overlapping sustainability and climate disclosures. *Nordea*. https://www.nordea.com/en/news/gri-sasb-cdp-making-sense-of-overlapping-sustainability-and-climate-disclosures

Garst, J., Maas, K., & Suijs, J. (2022). Materiality Assessment Is an Art, Not a Science: Selecting ESG Topics for Sustainability Reports. *California Management Review*, *65*(1), 64–90. https://doi.org/10.1177/00081256221120692

George, G., Osinga, E., Lavie, D., & Scott, B. (2016). Big Data and Data Science Methods for Management Research. *Academy of Management Journal*, *59*(5), 1493–1507. https://doi.org/10.5465/AMJ.2016.4005

Gerber, R., Smit, A., & Botha, M. (2024). An evaluation of environmental, social, and governance reporting in the agricultural sector. *Business Strategy & Development*, *7*(1), e316. https://doi.org/10.1002/bsd2.316

Gerring, J. (2012). *Social science methodology: A unified framework* (2nd ed). Cambridge University Press.

Goloshchapova, I., Poon, S.-H., Pritchard, M., & Reed, P. (2019). Corporate social responsibility reports: Topic analysis and big data approach. *The European Journal of Finance*, *25*(17), 1637–1654. https://doi.org/10.1080/1351847X.2019.1572637

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press. https://books.google.ca/books?hl=en&lr=&id=omivDQAAQBAJ&oi=fnd&pg=PR5&dq=Goodfellow,+I.,+et+al.+(2016).+%22Deep+Learning.%22+MIT+Press.&ots=MON6insFVW&sig=PvG2nhc8sCFEOrdRO-0KDG9nuOs

Goswami, K., Islam, M. K. S., & Evers, W. (2023). A Case Study on the Blended Reporting Phenomenon: A Comparative Analysis of Voluntary Reporting Frameworks and Standards—GRI, IR, SASB, and CDP. *The International Journal of Sustainability Policy and Practice*, *19*(2), 35.

GRI. (2024). GRI. https://www.globalreporting.org/

GRI and SASB. (2021, April 4). A Practical Guide to Sustainability Reporting Using GRI and SASB Standards. *SASB*. https://sasb.ifrs.org/knowledge-hub/practical-guide-to-sustainability-reporting-using-gri-and-sasb-standards/

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv:2203.05794*. https://doi.org/10.48550/arXiv.2203.05794

Gutierrez-Bustamante, M., & Espinosa-Leal. (2022). Natural Language Processing Methods for Scoring Sustainability Reports—A Study of Nordic Listed Companies. *Sustainability, MDPI*, *14*(15), 1–26.

Hadro, D., Fijałkowska, J., Daszyńska-Żygadło, K., Zumente, I., & Mjakuškina, S. (2022). What do stakeholders in the construction industry look for in non-financial disclosure and what do they get? *Meditari Accountancy Research*, *30*(3), 762–785. https://doi.org/10.1108/MEDAR-11-2020-1093/FULL/XML

Hahn, R., & Kühnen, M. (2013). Determinants of sustainability reporting: A review of results, trends, theory, and opportunities in an expanding field of research. *Journal of Cleaner Production*, *59*, 5–21. https://doi.org/10.1016/j.jclepro.2013.07.005

Hajaya, K. A. (2023). Adherence level to materiality in stand-alone GRI sustainability reporting in Arab nations: Does listing status matter. *Afro-Asian J. of Finance and Accounting*, *13*(3), 369–398. https://doi.org/10.1504/AAJFA.2023.132205

Hamza, S., & Jarboui, A. (2022). CSR or social impression management? Tone management in CSR reports. *Journal of Financial Reporting & Accounting*, *20*(3/4), 599–617. https://doi.org/10.1108/JFRA-04-2020-0115

Harjoto, M. A., Laksmana, I., & Lee, W. E. (2020). Female leadership in corporate social responsibility reporting: Effects on writing, readability and future social performance. *Advances in Accounting*, *49*, 100475. https://doi.org/10.1016/J.ADIAC.2020.100475

Harymawan, I., Ardianto, Ningtyas Soeprajitno, R. R. W., Ratri, M. C., & Paramitasari, Y. I. (2020). Text mining on sustainability reporting: A case study. *Journal of Security and Sustainability Issues*, *9*(May), 48–55. https://doi.org/10.9770/JSSI.2020.9.M(4)

Henriques, R., Gaio, C., & Costa, M. (2022). Sustainability Reporting Quality and Stakeholder Engagement Assessment: The Case of the Paper Sector at the Iberian Level. *Sustainability (Switzerland)*, *14*(21), 1–14. https://doi.org/10.3390/su142114404

Ho, J. C., Chen, T. H., & Wu, J. J. (2022). Are corporate social responsibility reports informative? Evidence from textual analysis of banks in China. *China Finance Review International*, *12*(1), 101–120. https://doi.org/10.1108/CFRI-04-2021-0081/FULL/XML

Hoepner, A. G. F., Dimatteo, S., Schaul, J., Yu, P.-S., & Musolesi, M. (2017). *Tweeting about Sustainability: Can Emotional Nowcasting Discourage Greenwashing?* (SSRN Scholarly Paper 2924088). https://doi.org/10.2139/ssrn.2924088

Hummel, K., & Jobst, D. (2024). An Overview of Corporate Sustainability Reporting Legislation in the European Union. *Accounting in Europe*, *0*(0), 1–36. https://doi.org/10.1080/17449480.2024.2312145

Hummel, K., Mittelbach-Hörmanseder, S., Cho, C. H., & Matten, D. (2022). Corporate social responsibility disclosure: A topic-based approach. *Accounting and Business Research*. https://doi.org/10.1080/00014788.2022.2071199

Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225. https://ojs.aaai.org/index.php/icwsm/article/view/14550

IFC. (2024). *Understanding the Global Reporting Frameworks*. International Finance Corporation (IFC), World Bank Group. https://www.ifcbeyondthebalancesheet.org/understanding-global-reporting-frameworks

Ignatov, K. (2021). Sentimental Sustainability: Does What Companies Say Tell More Than What Companies Do? *Financial Markets, Institutions & Instruments.* https://papers.ssrn.com/abstract=3933185

Ioannou, I., & Serafeim, G. (2017). The consequences of mandatory corporate sustainability reporting. *Harvard Business School Research Working Paper*, *11–100*. https://papers.ssrn.com/SOL3/PAPERS.CFM?ABSTRACT_ID=1799589

Jackson, G., Bartosch, J., Avetisyan, E., Kinderman, D., & Knudsen, J. S. (2020). Mandatory Non-financial Disclosure and Its Influence on CSR: An International Comparison. *Journal of Business Ethics*, *162*(2), 323–342. https://doi.org/10.1007/s10551-019-04200-0

Jacobs, B. W., & Singhal, V. R. (2017). The effect of the Rana Plaza disaster on shareholder wealth of retailers: Implications for sourcing strategies and supply chain governance. *Journal of Operations Management*, *49–51*(1), 52–66. https://doi.org/10.1016/j.jom.2017.01.002

Jafari, S. Q., Shokouhyar, S., & Shokoohyar, S. (2022). Producer-consumer sustainability continuum: Mutual understanding to implement extended producer responsibility. *Journal of Cleaner Production*, *374*, 133880. https://doi.org/10.1016/J.JCLEPRO.2022.133880

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: With Applications in R*. Springer US. https://doi.org/10.1007/978-1-0716-1418-1

Jaworska, S., & Nanda, A. (2018). Doing Well by Talking Good: A Topic Modelling-Assisted Discourse Study of Corporate Social Responsibility. *Applied Linguistics*, *39*(3), 373–399. https://doi.org/10.1093/APPLIN/AMW014

Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, *32*(3), 241–254.

Jørgensen, S., Mjøs, A., & Pedersen, L. J. T. (2022). Sustainability reporting and approaches to materiality: Tensions and potential resolutions. *Sustainability Accounting, Management and Policy Journal*, *13*(2), 341–361. https://doi.org/10.1108/SAMPJ-01-2021-0009

Jurafsky, D. (2000). *Speech and Language Processing*. Pearson Education India. https://www.thuvienso.net/upload/library/204335/preview/ebook-speech-and-language-processing-an-introduction-to-natural-language-processing-computational-linguistics-and-speech-recognition-part-1-daniel-jur.pdf

Kalra, V., & Aggarwal, R. (2017). Importance of Text Data Preprocessing & Implementation in RapidMiner. *ICITKM*, *14*, 71–75.

Kang, E., & Lam, N. B. (2023). The impact of environmental disclosure on initial public offering underpricing: Sustainable development in Singapore. *Corporate Social Responsibility and Environmental Management*, *30*(1), 119–133. https://doi.org/10.1002/csr.2342

Kang, H., & Kim, J. (2022). Analyzing and visualizing text information in corporate sustainability reports using natural language processing methods. *Applied Sciences*, *12*(11), 5614.

Kang, S. (2021). K-Nearest Neighbor Learning with Graph Neural Networks. *Mathematics*, *9*(8), 830. https://doi.org/10.3390/MATH9080830

Karnani, A. (2010). The case against corporate social responsibility. *Wall Street Journal*, *23*(14), 1–5.

Kayser, S. (2016). The rise and role of the 'Accord' and the 'Alliance' in response to the collapse of Rana Plaza: Real or symbolic change? *ERB*. https://erb.umich.edu/2016/04/18/the-rise-and-role-of-the-accord-and-the-alliance-in-response-to-the-collapse-of-rana-plaza-real-or-symbolic-change/

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review*, *91*(6), 1697–1724.

Kim, E. H., & Lyon, T. P. (2015). Greenwash vs. Brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. *Organization Science*, *26*(3), 705–723. https://doi.org/10.1287/ORSC.2014.0949/ASSET/IMAGES/LARGE/ORSC.2014.0949-F01.JPEG

Kiriu, T., & Nozaki, M. (2020a). A Text Mining Model to Evaluate Firms' ESG Activities: An Application for Japanese Firms. *Asia-Pacific Financial Markets*, *27*(4), 621–632. https://doi.org/10.1007/s10690-020-09309-1

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—A systematic literature review. *Information and Software Technology*, *51*(1), 7–15. https://doi.org/10.1016/j.infsof.2008.09.009

Klimczak, K. M., Hadro, D., & Meyer, M. (2023). Executive communication with stakeholders on sustainability: The case of Poland. *Accounting in Europe*. https://doi.org/10.1080/17449480.2023.2213242

Knight, W. (2021). *The Left and the Right Speak Different Languages—Literally*. https://www.wired.com/story/left-right-speak-different-languages-literally/

Kountouri, I., Manousakis, E., & Tsekrekos, A. E. (2019). Latent semantic analysis of corporate social responsibility reports (with an application to Hellenic firms). *International Journal of Disclosure and Governance*, *16*(1), 1–19. https://doi.org/10.1057/s41310-018-0053-z

Kozlowski, A., Searcy, C., & Bardecki, M. (2015). Corporate sustainability reporting in the apparel industry an analysis of indicators disclosed. *International Journal of Productivity and Performance Management*, *64*(3), 377–397. https://doi.org/10.1108/IJPPM-10-2014-0152/FULL/XML

KPMG international. (2014). *The essentials of materiality assessment—KPMG China*. KPMG. https://home.kpmg/cn/en/home/insights/2014/10/materiality-assessment.html

KPMG International. (2022). *Big shifts, small steps> Survey of Sustainability Reporting 2022*. https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2023/04/big-shifts-small-steps.pdf

Krippendorff, K. (2019). *Content Analysis: An Introduction to Its Methodology*. SAGE Publications, Inc. https://doi.org/10.4135/9781071878781

Kvasničková Stanislavská, L., Pilař, L., Fridrich, M., Kvasnička, R., Pilařová, L., Afsar, B., & Gorton, M. (2023). Sustainability reports: Differences between developing and developed countries. *Frontiers in Environmental Science*, *11*. https://doi.org/10.3389/FENVS.2023.1085936/FULL

Labowitz, S., & Baumann-Pauly, D. (2014). Business as usual is not an option. *New York University Center for Business and Human Rights*. https://www.stern.nyu.edu/sites/default/files/assets/documents/con_047408.pdf

Lai, A., & Stacchezzini, R. (2021). Organisational and professional challenges amid the evolution of sustainability reporting: A theoretical framework and an agenda for future research. *Meditari Accountancy Research*, *29*(3), 405–429. https://doi.org/10.1108/MEDAR-02-2021-1199

Laufer, W. S. (2003). Social Accountability and Corporate Greenwashing. *Journal of Business Ethics*, *43*(3), 253–261. https://doi.org/10.1023/A:1022962719299

Leavitt, K., Schabram, K., Hariharan, P., & Barnes, C. M. (2021). Ghost in the machine: On organizational theory in the age of machine learning. *Academy of Management Review*, *46*(4), 750–777.

Lewis, C., & Young, S. (2019). Fad or future? Automated analysis of financial text and its implications for corporate reporting. *Accounting and Business Research*, *49*(5), 587–615. https://doi.org/10.1080/00014788.2019.1611730

Li, C., & Huang, M. (2023). *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-023-30301-5

Li, M., & Zhao, L. (2021). Exploring Global Fashion Sustainability Practices through Dictionary-Based Text Mining. *Clothing and Apparels Research Journal*, 1–16. https://doi.org/10.1177/0887302X21998268

Li, Y., & Rockinger, M. (2024). Unfolding the Transitions in Sustainability Reporting. *Sustainability*, *16*(2), Article 2. https://doi.org/10.3390/su16020809

Liang, X., & Wu, H. (2022). Does the Tone in Corporate Social Responsibility Reports Misdirect Analysts' Forecasts in China? *Sustainability*, *14*(24), 16631. https://doi.org/10.3390/SU142416631

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22.

Liew, W. T., Adhitya, A., & Srinivasan, R. (2014). Sustainability trends in the process industries: A text mining-based analysis. *Computers in Industry*, *65*(3), 393–400. https://doi.org/10.1016/j.compind.2014.01.004

Lindgren, C., Huq, A. M., & Carling, K. (2021a). Who Are the Intended Users of CSR Reports? Insights from a Data-Driven Approach. *Sustainability*, *13*(3), 1070. https://doi.org/10.3390/su13031070

Liu, J., Sun, J., & Fang, M. (2022). Construction and Analysis of Corporate Social Responsibility Report and Equity Capital Cost Algorithm Model Based on Intelligent Computing. *Security and Communication Networks*, *2022*. https://doi.org/10.1155/2022/3306735

Liu, S.-H., Chen, S.-Y., & Li, S.-T. (2017). Text-Mining Application on CSR Report Analytics: A Study of Petrochemical Industry. *6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, 76–81. https://doi.org/10.1109/IIAI-AAI.2017.164

Locke, R., Amengual, M., & Mangla, A. (2009). Virtue out of Necessity? Compliance, Commitment, and the Improvement of Labor Conditions in Global Supply Chains. *Politics & Society*, *37*(3), 319–351. https://doi.org/10.1177/0032329209338922

Lu, J., & Jagoda, K. (2023). Sentiment analysis in sustainability accounting reporting: Does the tone reveal future environmental performance? *International Journal of Accounting and Finance*. https://www.inderscienceonline.com/doi/10.1504/IJAF.2023.134518

Luo, X., Wang, H., Raithel, S., & Zheng, Q. (2015). Corporate social performance, analyst stock recommendations, and firm future returns. *Strategic Management Journal*, *36*(1), 123–136. https://doi.org/10.1002/SMJ.2219

Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate Environmental Disclosure under Threat of Audit. *Journal of Economics & Management Strategy*, *20*(1), 3–41. https://doi.org/10.1111/J.1530-9134.2010.00282.X

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, *1*(14), 281–297.

Mangsor, N. S. M. N., Nasir, S. A. M., Yaacob, W. F. W., Ismail, Z., & Rahman, S. A. (2022). Analysing corporate social responsibility reports using document clustering and topic modeling techniques. *Indonesian Journal of Electrical Engineering and Computer Science*, *26*(3), 1546–1555. https://doi.org/10.11591/IJEECS.V26.I3.PP1546-1555

Mani, V., Gunasekaran, A., & Delgado, C. (2018). Supply chain social sustainability: Standard adoption practices in Portuguese manufacturing firms. *International Journal of Production Economics*, *198*, 149–164. https://doi.org/10.1016/J.IJPE.2018.01.032

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press.

Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, Norms, and Selective Disclosure: A Global Study of Greenwashing. *Organization Science*, *27*(2), 483–504. https://doi.org/10.1287/orsc.2015.1039

Maulida, S., & Hakim, B. A. (2022). Twitter Sentiment Analysis on Green Finance. *Accounting and Sustainability*, *1*(1), Article 1. https://doi.org/10.58968/as.v1i1.256

Merkl-Davies, D. M., & Brennan, N. M. (2008). *Discretionary Disclosure Strategies in Corporate Narratives: Incremental Information or Impression Management?* (SSRN Scholarly Paper 1089447). https://papers.ssrn.com/abstract=1089447

Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. *Critical Perspectives on Accounting*, *33*, 59–78. https://doi.org/10.1016/j.cpa.2014.10.003

Michelon, G., Trojanowski, G., & Sealy, R. (2022). Narrative Reporting: State of the Art and Future Challenges. *Accounting in Europe*, *19*(1), 7–47. https://doi.org/10.1080/17449480.2021.1900582

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space* (arXiv:1301.3781). arXiv. http://arxiv.org/abs/1301.3781

Miller, H. G., & Mork, P. (2013). From data to decisions: A value chain for big data. *It Professional*, *15*(1), 57–59.

Miner, G. D., Elder, J., Fast, A., Hill, T., Nisbet, R., & Delen, D. (2012). *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications*. Academic Press.

Modapothala, J. R., & Issac, B. (2009). Angin. *Software Technology and Engineering*, 151–155. https://doi.org/10.1142/9789814289986_0028

Mohan, V., Vijayarani, S., & Janani, M. R. (2016). Advanced Computational Intelligence. *An International Journal (ACII)*, *3*(1). https://doi.org/10.5121/acii.2016.3104

Moreno, A. I., & Caminero, T. (2022). Application of text mining to the analysis of climate-related disclosures. *International Review of Financial Analysis*, *83*, 102307. https://doi.org/10.1016/J.IRFA.2022.102307

Mućko, P. (2021). Sentiment analysis of CSR disclosures in annual reports of EU companies. *Procedia Computer Science*, *192*, 3351–3359. https://doi.org/10.1016/j.procs.2021.09.108

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association : JAMIA*, *18*(5), 544–551. https://doi.org/10.1136/amiajnl-2011-000464

Nast, C. (2023, April 24). *10 years since Rana Plaza, not enough has changed*. Vogue Business. https://www.voguebusiness.com/sustainability/10-years-since-rana-plaza-not-enough-has-changed

Ngu, S. B., & Amran, A. (2021). Materiality disclosure in sustainability reporting: Evidence from Malaysia. *Asian Journal of Business and Accounting*, *14*(1), 225–252. https://doi.org/10.22452/ajba.vol14no1.9

Ning, X., Yim, D., & Khuntia, J. (2021). Online Sustainability Reporting and Firm Performance: Lessons Learned from Text Mining. *Sustainability*, *13*(3), 1069. https://doi.org/10.3390/su13031069

Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management*, *53*, 102104. https://doi.org/10.1016/j.ijinfomgt.2020.102104

Niveditha, R., Ns, N. K., Parimi, M. R., Aaditya Raam, S., & Babu, S. (2020). Develop CSR Themes using Text-Mining and Topic Modelling Techniques. *Proceedings - 2020 IEEE International Conference on Cloud Computing in Emerging Markets, CCEM 2020*, 67–71. https://doi.org/10.1109/CCEM50674.2020.00024

Ortar, L. (2018). Materiality matrixes in sustainability reporting: An empirical examination. *Available at SSRN 3117749*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3117749

Padilla-Garrido, N., Aguado-Correa, F., Rabadán-Martín, I., López-Jiménez, J. M., De La Vega-Jiménez, J. J., & Peletier-Ribera, I. (2024). Materiality analysis in sustainability reporting: Insights from large Spanish companies. *Corporate Social Responsibility and Environmental Management*, csr.2866. https://doi.org/10.1002/csr.2866

Pan, Y. (2016). Sustainability trends in financial services sector: Evidence from Europe and North America. *2016 13th International Conference on Service Systems and Service Management, ICSSSM 2016*. https://doi.org/10.1109/ICSSSM.2016.7538549

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1–135.

Pedhazur, E. J., & Schmelkin, L. P. (2013). *Measurement, Design, and Analysis: An Integrated Approach*. Psychology Press.

Picard, Ohl, & Beagent. (2023). *Why impact matters for company valuations*. PwC. https://www.pwc.com/gx/en/services/audit-assurance/corporate-reporting/esg-reporting/esg-impact-company-valuation.html

Pimonenko, T., Bilan, Y., Horák, J., Starchenko, L., & Gajda, W. (2020). Green Brand of Companies and Greenwashing under Sustainable Development Goals. *Sustainability*, *12*(4), Article 4. https://doi.org/10.3390/su12041679

Pizzi, S., Principale, S., & de Nuccio, E. (2022). Material sustainability information and reporting standards. Exploring the differences between GRI and SASB. *Meditari Accountancy Research*, *31*(6), 1654–1674. https://doi.org/10.1108/MEDAR-11-2021-1486

Pollach, I. (2018). Issue Cycles in Corporate Sustainability Reporting: A Longitudinal. *Environmental Communication*, *12*(2), 247–260. https://doi.org/10.1080/17524032.2016.1205645

Raghupathi, V., Ren, J., & Raghupathi, W. (2020). Identifying Corporate Sustainability Issues by Analyzing Shareholder Resolutions: A Machine-Learning Text Analytics Approach. *Sustainability, MDPI*, *12*(11), 4753. https://doi.org/10.3390/SU12114753

Ramakrishnan, J., Liu, T., Zhang, F., Seshadri, K., Yu, R., & Gou, Z. (2023a). A decision tree-based modeling approach for evaluating the green performance of airport buildings. *Environmental Impact Assessment Review*, *100*, 107070. https://doi.org/10.1016/J.EIAR.2023.107070

Raman, N., Bang, G., & Nourbakhsh, A. (2020). Mapping ESG trends by distant supervision of neural language models. *Machine Learning and Knowledge Extraction*, *4*, 453–468. https://doi.org/10.3390/make2040025

Reinecke, J., & Donaghey, J. (2015). After Rana Plaza: Building coalitional power for labour rights between unions and (consumption-based) social movement organisations. *Organization (London, England)*, *22*(5), 720–740. https://doi.org/10.1177/1350508415585028

Resnik, S., & Koklič, M. K. (2018). User-generated tweets about global green brands: A sentiment analysis approach. *Trziste= Market*, *30*(2), 125–145.

Reuter, N., Debortoli, S., & Müller, O. (2014). Identifying The Role Of Energy-Related Environmental Sustainability Using Text Mining. *Ecis*, 0–15.

Rhoden, I., Ball, C. S., Vögele, S., & Kuckshinrichs, W. (2023). Minding the gap-relating disclosure to contexts of sustainability reporting in the automotive industry. *Corporate Social Responsibility and Environmental Management*, *30*(2), 846–857. https://doi.org/10.1002/CSR.2392

Riani, R., & Rusydiana, A. S. (2023). Exploring Sentiment Analysis of Sustainable Finance Initiatives: A Text Mining Approach. *Accounting and Sustainability*, *2*(1), Article 1. https://doi.org/10.58968/as.v2i1.266

Ribeiro, F. N., Araújo, M., Gonçalves, P., André Gonçalves, M., & Benevenuto, F. (2016). SentiBench—A benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science*, *5*(1), 23. https://doi.org/10.1140/epjds/s13688-016-0085-1

Roca, L. C., & Searcy, C. (2012). An analysis of indicators disclosed in corporate sustainability reports. *Journal of Cleaner Production*, *20*(1), 103–118. https://doi.org/10.1016/J.JCLEPRO.2011.08.002

Rocca, L., Giacomini, D., & Zola, P. (2020). Environmental disclosure and sentiment analysis: State of the art and opportunities for public-sector organisations. *Meditari Accountancy Research*, *29*(3), 617–646. https://doi.org/10.1108/MEDAR-09-2019-0563/FULL/HTML

Saha, A., & Nabareseh, S. (2015). Communicating corporate social responsibilities: Using text mining for a comparative analysis of banks in India and Ghana. *Mediterranean Journal of Social Sciences*, *6*(3), 11–20. https://doi.org/10.5901/mjss.2015.v6n3s1p11

Sai, P. K., Gupta, P., & Fernandes, S. F. (2019). *Analysing Performance of Company through Annual reports using Text Analytics*. 21–31. https://doi.org/10.1109/ICD47981.2019.9105673

Sambhanthan, A., & Potdar, V. (2015). Green Business Practices for Software Development Companies: An Explorative Text Analysis of Business Sustainability Reports. *International Journal of Enterprise Information Systems*, *11*(3), 13–26. https://doi.org/10.4018/IJEIS.2015070102

Sambhanthan, A., & Potdar, V. (2016). Waste management strategies for Software Development companies: An explorative text analysis of business sustainability reports. *2016 IEEE 14th International Conference on Software Engineering Research, Management and Applications (SERA)*, 179–184. https://doi.org/10.1109/SERA.2016.7516144

Samkin, G. (2012). Changes in sustainability reporting by an African defence contractor: A longitudinal analysis. *Meditari Accountancy Research*, *20*(2), 134–166. https://doi.org/10.1108/10222521211277834/FULL/XML

SASB. (2021). SASB. https://sasb.org/

Shahi, A. M., Issac, B., & Modapothala, J. R. (2012). Enhanced intelligent text categorization using concise keyword analysis. *ICIMTR 2012 - 2012 International Conference on Innovation, Management and Technology Research*, 574–579. https://doi.org/10.1109/ICIMTR.2012.6236461

Shahi, A. M., Issac, B., & Modapothala, J. R. (2014). Automatic analysis of corporate sustainability reports and intelligent scoring. *International Journal of Computational Intelligence and Applications*, *13*(1). https://doi.org/10.1142/S1469026814500060

Sharma, A., Kaushik, K., Awasthy, P., & Gawande, A. (2022). Leveraging Text Mining for Trend Analysis and Comparison of Sustainability Reports: Evidence from Fortune 500 Companies. *American Business Review*, *25*(2), 8. https://doi.org/10.37625/abr.25.2.416-438

Slacik, J., & Greiling, D. (2020). Compliance with materiality in G4-sustainability reports by electric utilities. *International Journal of Energy Sector Management*, *14*(3), 583–608.

Smeuninx, N., De Clerck, B., & Aerts, W. (2016). Measuring the Readability of Sustainability Reports: A Corpus-Based Analysis Through Standard Formulae and NLP. *International Journal of Business Communication*, *57*(1), 52–58. https://doi.org/10.1177/2329488416675456

Son, B.-G. (2005). WordStat 5.0: Content analysis software offers operations researchers valuable tool for exploring vast, untapped amounts of textual data. *OR/MS Today*, *32*(5), 54–59.

Stanik, C., Pietz, T., & Maalej, W. (2021). Unsupervised Topic Discovery in User Comments. *2021 IEEE 29th International Requirements Engineering Conference (RE)*, 150–161. https://doi.org/10.1109/RE51729.2021.00021

Suddaby, R., & Panwar, R. (2022). On the Complexity of Managing Transparency. *California Management Review*, *65*(1), 5–18. https://doi.org/10.1177/00081256221128766

Sustainability Knowledge Group. (2021). GRI and SASB Reporting Standards: Working Together. *Sustainability Knowledge Group*. https://sustainabilityknowledgegroup.com/gri-and-sasb-reporting-standards-working-together/

Szekely, N., & Vom Brocke, J. (2017). What can we learn from corporate sustainability reporting? Deriving propositions for research and practice from over 9,500 corporate sustainability reports published between 1999 and 2015 using topic modelling technique. *PLOS ONE*, *12*(4), e0174807. https://doi.org/10.1371/JOURNAL.PONE.0174807

Tate, W., Ellram, L., & Kirchoff, J. (2009). Corporate social responsibility reports: A thematic analysis related to supply chain management. *Journal of Supply Chain Management*, *46*(1), 19–44. https://doi.org/10.1111/j.1745-493X.2009.03184.x

Thangam, D., Chavadi, C., & Manoharan, S. (n.d.). *Corporate Sustainability, Greenwash, and Its Impact on Environment, Society, and Governance*. Retrieved August 30, 2024, from https://www.researchgate.net/publication/381584492_CORPORATE_SUSTAINABILITY_GREENWASH_AND_ITS_IMPACT_ON_ENVIRONMENT_SOCIETY_AND_GOVERNANCE_CSG-ESG_2024

Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, *63*(1), 163–173. https://doi.org/10.1002/asi.21662

Thelwell, K. (2021, May 26). The Impact of Fast Fashion in Bangladesh. *The Borgen Project*. https://borgenproject.org/fast-fashion-in-bangladesh/

Thomas, D. A. (2014). Searching for Significance in Unstructured Data: Text Mining with Leximancer. *European Educational Research Journal*, *13*(2), 235–256. https://doi.org/10.2304/eerj.2014.13.2.235

Tiscini, R., Martiniello, L., & Lombardi, R. (2021). Circular economy and environmental disclosure in sustainability reports: Empirical evidence in cosmetic companies. *Business Strategy and Environment*, *31*(3), 892–907. https://doi.org/10.1002/bse.2924

Tomlinson, B., & Godshall, L. (2023). *ESG regulation gives materiality a bigger stake*. EY. https://www.ey.com/en_us/insights/assurance/esg-regulation-gives-materiality-a-bigger-stake

Torelli, R., Balluchi, F., & Lazzini, A. (2020). Greenwashing and environmental communication: Effects on stakeholders' perceptions. *Business Strategy and the Environment*, *29*(2), 407–421. https://doi.org/10.1002/bse.2373

Tóth, Á., & Suta, A. (2021). Global sustainability reporting in the automotive industry via the eXtensible business reporting language. *Chemical Engineering Transactions*, *88*, 1087–1092. https://doi.org/10.3303/CET2188181

Tóth, Á., Suta, A., & Szauter, F. (2022). Interrelation between the climate-related sustainability and the financial reporting disclosures of the European automotive industry. *Clean Technologies and Environmental Policy*, *24*(1), 437–445. https://doi.org/10.1007/S10098-021-02108-W/TABLES/4

Tóth, Á., Szigeti, C., & Suta, A. (2021). Carbon Accounting Measurement with Digital Non-Financial Corporate Reporting and a Comparison to European Automotive Companies Statements. *Energies*, *14*(18), 5607. https://doi.org/10.3390/en14185607

Tremblay, M. C., Mason, R. A., & Forsgren, N. (2021). Theories in Flux: Reimagining Theory Building in the Age of Machine Learning. *MIS Quarterly*, *45*(1), 455–459. https://doi.org/10.25300/MISQ/2021/15434.1.1

Tremblay, M. C., Parra, C. M., & Castellanos, A. (2015). Corporate social responsibility reports: Understanding topics via text mining. *2015 Americas Conference on Information Systems, AMCIS 2015*.

Uyar, A., Koseoglu, M. A., Kılıç, M., & Mehraliyev, F. (2021). Thematic structure of sustainability reports of the hospitality and tourism sector: A periodical, regional, and format-based analysis. *Current Issues in Tourism*, *24*(18), 2602–2627. https://doi.org/10.1080/13683500.2020.1847050

Vaio, A. D., Hassan, R., D'Amore, G., & Strologo, A. D. (2022). Digital Technologies for Sustainable Waste Management On-Board Ships: An Analysis of Best Practices From the Cruise Industry. *IEEE Transactions on Engineering Management*. https://doi.org/10.1109/TEM.2022.3197241

Vieira, I. L., Silva, E. R. da, Martini, L. C. de, & Rangel, L. A. D. (2021). Proposal for an analytical model of materiality in the Global Reporting Initiative Standards reports. *Production*, *31*, e20210003.

Vigneau, L., & Adams, C. A. (2023). The failure of transparency as self-regulation. *Sustainability Accounting, Management and Policy Journal*, *14*(4), 852–876.

Vouros, P., Nomikos, S., Halkos, G., Evangelinos, K., Sfakianaki, E., Konstandakopoulou, F., Fotiadis, S., Karagiannis, I., Skouloudis, A., & Nikolaou, I. E. (2020). Introducing fundamental accountability principles in sustainability reporting assessment: A cross-sectoral analysis from the Greek business sector. *Environmental Quality Management*, *29*(4), 33–50. https://doi.org/10.1002/tqem.21689

Wallbaum, H., Krank, S., & Teloh, R. (2011). Prioritizing Sustainability Criteria in Urban Planning Processes: Methodology Application. *Journal of Urban Planning and Development*, *137*(1), 20–28. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000038

Wang, W., Kang, W., & Mu, J. (2023). *Mapping research to the sustainable development goals (sdgs)*. https://www.researchsquare.com/article/rs-2544385/latest

Wang, X., Yuen, K. F., Wong, Y. D., & Li, K. X. (2020). How can the maritime industry meet Sustainable Development Goals? An analysis of sustainability reports from the social entrepreneurship perspective. *Transportation Research Part D: Transport and Environment*, *78*, 102173. https://doi.org/10.1016/J.TRD.2019.11.002

Wanner, J., & Janiesch, C. (2019). Big data analytics in sustainability reports: An analysis based on the perceived credibility of corporate published information. *Business Research*, *12*(1), 143–173. https://doi.org/10.1007/s40685-019-0088-4

Whittingham, K. L., Earle, A. G., Argiolas, A., & Hiz, D. I. L. la. (2022). The impact of the United Nations Sustainable Development Goals on corporate sustainability reporting. *BRQ Business Research Quarterly*, *26*(1), 45–61. https://doi.org/10.1177/23409444221085585

Wu, P.-J., & Huang, P.-C. (2018). Business analytics for systematically investigating sustainable food supply chains. *Journal of Cleaner Production*, *203*, 968–976.

Wu, S. R., Shao, C., & Chen, J. (2018). Approaches on the Screening Methods for Materiality in Sustainability Reporting. *Sustainability, MDPI*, *10*(9), 3233. https://doi.org/10.3390/SU10093233

Yang, B. M., & Yang, O. S. (2022). Assessing the Effect of Dynamic Capabilities on the ESG Reporting and Corporate Performance Relationship With Topic Modeling: Evidence From Global Companies. *Frontiers in Psychology*, *13*, 898935. https://doi.org/10.3389/FPSYG.2022.898935/BIBTEX

Ye, M., Lu, W., Flanagan, R., & Chau, K. W. (2019). Corporate social responsibility "glocalisation": Evidence from the international construction business. *Corporate Social Responsibility and Environmental Management*, *27*(2), 655–669. https://doi.org/10.1002/csr.1831

Ye, M., Lu, W., & Xue, F. (2022). Impact of Institutional Distance on Environmental and Social Practices in Host Countries: Evidence from International Construction Companies. *Journal of Construction Engineering and Management*, *148*(1). https://doi.org/10.1061/(ASCE)CO.1943-7862.0002226

Yu, E. P., Luu, B. V., & Chen, C. H. (2020). Greenwashing in environmental, social and governance disclosures. *Research in International Business and Finance*, *52*, 101192. https://doi.org/10.1016/j.ribaf.2020.101192

Zakaria, M., Aoun, C., & Liginlal, D. (2021). Objective sustainability assessment in the digital economy: An information entropy measure of transparency in corporate sustainability reporting. *Sustainability, MDPI*, *13*(3), 1–18. https://doi.org/10.3390/su13031054

Zharfpeykan, R., & Akroyd, C. (2023). Evaluating the outcome effectiveness of the global reporting initiative transitions. *Sustainability Accounting, Management and Policy Journal*, *14*(6), 1101–1125.

Zhou, A. (2021). Communicating corporate LGBTQ advocacy: A computational comparison of the global CSR discourse. *Public Relations Review*, *47*(4), 102061. https://doi.org/10.1016/J.PUBREV.2021.102061

Zhou, K., Jin, X., Li, X., & Tao, Y. (2023). Enhancing sustainable development through effective disclosure: Corporate environmental performance and readability. *Corporate Social Responsibility and Environmental Management*. https://doi.org/10.1002/CSR.2569

Zhou, Y., Li, X., Wang, X., & Yuen, K. F. (2022). Intelligent container shipping sustainability disclosure via stakeholder sentiment views on social media. *Marine Policy*, *135*. https://doi.org/10.1016/j.marpol.2021.104853

Zhou, Y., Wang, X., & Yuen, K. F. (2021). Sustainability disclosure for container shipping: A text-mining approach. *Transport Policy*, *110*, 465–477. https://doi.org/10.1016/j.tranpol.2021.06.020

Chapter 2: Investigating "Transparency" in Sustainability Reports of the Apparel Industry Using Text Mining and Natural Language Processing

Mahsa Mohammadrezaei

Telfer School of Management, University of Ottawa, Ottawa, Canada

PhD Thesis

University of Ottawa, Ottawa, Canada

Fall 2024

1 Introduction

In recent years, there has been a dramatic increase in corporate disclosure of non-financial data related to social and environmental issues (Lindgren et al., 2021). A wide variety of terms have been employed for this type of reporting, including "sustainability," "accountability," “sustainable development," "corporate social responsibility," "corporate responsibility," “non-financial disclosure,” "corporate citizenship," and "triple bottom line" reporting (M. Li & Zhao, 2021; Shahi et al., 2014). In this study, we use the term “SR,” the acronym for “sustainability report,” to cover all these terms and any similar type of reporting.

Despite the increase in volume, concerns have been raised regarding the quality and precision of disclosures (Aras & Crowther, 2008; Billio et al., 2021; Boiral & Heras-Saizarbitoria, 2020; Bushee et al., 2018; Clark et al., 2019; Fabrizio & Kim, 2019; Kim & Lyon, 2015). The most frequent critiques highlight issues related to transparency, credibility, and relevance of information within SRs, indicating the need for improved reporting practices (Cho et al., 2015; Gray, 2010; Lock & Seele, 2016). In addition, the sheer volume, diversity, and complexity of data within SRs pose significant challenges for their evaluation, hindering analysis, interpretation, and comparison and prompting questions about the "transparency" of sustainability reporting, with concerns about irrelevant data disclosures and over-generalized content (Boiral & Henri, 2015; Dubbink et al., 2008).

This array of issues with non-financial reporting can be grouped into two key categories: 1) what is being reported, meaning the significance and relevance of the disclosed content (frequently referred to as "materiality in reporting"); 2) how this is being reported, meaning the objectivity of the information being provided, particularly whether both strength/opportunities and risks and areas for improvement are being disclosed (increasingly referred to as "balanced reporting").

Calls for a rigorous application of the materiality principle have emerged as a central component of sound SR (Diouf & Boiral, 2017; Eccles et al., 2012a). The materiality principle specifies that only what is considered relevant and salient to the firm and industry in which it operates warrants reporting and dictates that SRs should include only those elements that are most crucial for informed decision-making by their stakeholders (Calabrese et al., 2016). By prioritizing material aspects, companies can provide more pertinent disclosures, thereby enhancing the overall quality of their reports (Adams & Abhayawansa, 2022; Talbot & Boiral, 2018).

However, addressing materiality is only part of the equation. The way companies convey the information - "how" they report on both progress and challenges - is equally crucial. Good practice dictates that SRs should be as transparent as possible, communicating both opportunities and risks/areas for improvement of relevant topics (SASB, 2024). Referred to as "balanced reporting," this has been increasingly emphasized as key to transparent disclosure (Corporate ESG Reporting, 2024; GRI, 2024; Foley, 2024; Groothuis, 2019; KPMG, 2023; KPMG international, 2014).

In response to these challenges, numerous sustainability accounting frameworks, guidelines and best practices have been developed. Amongst these, two stand out for their widespread adoption. The Global Reporting Initiative (GRI) framework is a comprehensive, globally recognized standard that emphasizes broad ESG impacts, aiming to provide transparency and accountability across a wide range of stakeholder groups, including communities, employees, and policymakers (GRI, 2024). The Sustainability Accounting Standards Board (SASB) framework is a pivotal industry-specific framework focused on enhancing the material relevance and accuracy of corporate disclosures to financial stakeholders (e.g. investors) (Busco et al., 2020; Kuh et al., 2020). In addition to the reporting frameworks, audit and consulting firms have responded to these challenges by developing environmental, social, and governance (ESG) services that help firms navigate the non-financial disclosure landscape. Firms like KPMG, Deloitte, and PWC have been active in the promotion of materiality and balance as a means of fostering trust and credibility on par with financial reporting and thus enabling stakeholders to make well-informed decisions (Santos et al., 2023).

However, many companies have been criticized for publishing SRs that provide little value to stakeholders. From lack of clarity in reporting metrics and inconsistent data presentation to selective disclosure of positive outcomes while omitting negative impacts, these reports often fail to offer a comprehensive and accurate view of a company’s sustainability efforts (Liao and Shi, 2023). This has led to accusations that some companies are engaging in what has been commonly referred to as “greenwashing.” The latter involves the strategic manipulation of sustainability communications to create a misleading impression of environmental responsibility or ethical practices. These may mislead governments, consumers, investors, and the public by making unsubstantiated or exaggerated claims about ESG aspects of a product, service, or overall operations (Santos et al.,2023). Such practices undermine the integrity of sustainability communications and pose risks like reputational damage and increased regulatory scrutiny (Gil-Cordero et al., 2021; Hameed et al., 2021; Kim & Lyon, 2015; Montgomery et al., 2023; Santos et al., 2023). While it is widely recognized that SRs often use optimistic language to present a favourable image to stakeholders (Friske et al., 2023),

less understood is the extent of this positivity and how to assess and measure what might be considered "excessive optimism" or "exaggeration." This issue is critically important because overly optimistic portrayals and exaggeration of sustainability efforts or the downplaying of negative aspects can quickly deteriorate point of creating an unrealistic depiction of reality – a major issue for investors, civil society and regulators.

Despite the importance of addressing these issues, academic research on SR materiality and balance has often fallen short. Conventional content analysis methods, which have tended to be predominantly manual or based on simple text mining such as word frequency and topic modelling, are not capable of offering the depth and breadth necessary for the type of nuanced analysis required to address the problem (Sai et al., 2019, Reuter et al., 2014, Tremblay et al., 2015, Pan, 2016, Saha & Nabareseh, 2015, Li & Zhao, 2021). Depth, in this context, refers to the computational ability of a tool to manage and interpret complex data relationships and execute sophisticated analyses in a manner that generates considerable insight regarding the issue. Breadth, on the other hand, pertains to the capacity of a method to handle large datasets effectively. In terms of depth, these conventional methods can be understood as surface-level analyses as they primarily involve straightforward counting of words or phrases and rudimentary categorization of content. Such approaches are useful for initial examinations of large text volumes to pinpoint dominant themes or terms, but they are unable to capture the nuanced meanings and deeper implications inherent in sustainability reports. These approaches also exhibit limitations in terms of breadth as they are either not feasible on large datasets (as is the case for manual analysis) or are feasible but have been applied to a narrow dataset, likely because of a constricted data collection process (as is the case with much of the text mining that has been applied to SR analysis – see paper 1 in this thesis).

This study contributes to and advances this literature by providing a more granular and nuanced analysis of a relatively large volume of SRs over time. By systematically exploring materiality and balance at the company and industry levels over seven years, this research offers greater depth and breadth in SR analysis compared to both studies employing manual content analysis methods and studies that have incorporated advanced techniques (Angin et al., 2022; Ehrhardt & Nguyen, 2021; Y. Li & Rockinger, 2024).

Our research context is the SR of companies implicated in the Rana Plaza disaster, this tragic event highlighted the urgent need for improved sustainability practices and transparent reporting, leading to initiatives like the Accord on Fire and Building Safety in Bangladesh and the Alliance for Bangladesh Worker Safety (often referred to as the “Accord” and “Alliance”), which aim to enhance

safety and sustainability in the apparel sector. These initiatives, along with other related efforts, have increased the volume of SR disclosures and underscored the need for both material and balanced reporting, particularly on labour conditions (Bair et al., 2020; Baumann-Pauly et al., 2018; Donaghey & Reinecke, 2018; Kashyap, 2015; Labowitz & Baumann-Pauly, 2014). We use this setting as an "extreme case" context where SR reporting is suddenly much more abundant, and there is considerable pressure to be more transparent and forthcoming about the challenges that companies are facing. Companies that have signed initiatives like the Accord and Alliance have committed to improving their sustainability efforts and to report on these. Unfortunately, in the years following Rana Plaza, it has been well documented that there has been little if any progress and improvement in working conditions and that, as such, there seems to have been a significant and concerning underperformance when compared to the promises made immediately after the disaster (Barua et al., 2021).

However, the relationship between actual sustainability performance and corporate sustainability reporting (SR) among companies affected by the Rana Plaza disaster remains poorly understood. This research addresses this gap by examining SRs within the apparel industry, focusing on both materiality (theme extraction) and balance. Our study specifically analyzes the SRs of individual companies most impacted by the disaster (103 reports) and conducts an aggregate analysis of the broader industry (322 reports). This unique and timely research context provides valuable insights into how these companies communicate their sustainability efforts. To do so, we develop an analytical framework based on the GRI and SASB standards (GRI & SASB, 2021) that allow us to analyze and assess materiality and balance in an integrated fashion. This leads us to our central research question: Does the apparel industry effectively report on material topics, and if so, is the reporting balanced?

We employ a two-tiered approach, examining both individual companies and the industry as a whole. At the company level, we quantified alignment with predefined material topics and analyzed the sentiment of extracted sentences. This revealed which sustainability topics are most discussed and how companies balance addressing strengths and opportunities versus risks and areas for improvement. At the aggregate level, we applied a similar method, focusing on sentences with the strongest positive and negative sentiments across the industry. This approach allowed us to identify the most prominent topics and sentiments within the broader apparel sector, offering a comprehensive view of industry-wide communication strategies for sustainability.

Our analysis of SRs uncovers key insights into SR practices within the apparel industry. First, companies consistently focused on material topics identified through the GRI and SASB frameworks throughout the study period, indicating an established core of industry-relevant issues. However, the thematic alignment analysis revealed an unexpected emphasis on environmental issues over labour concerns, particularly among Accord and Alliance companies, despite the latter's heightened importance post-Rana Plaza. The analysis also identified substantial variability in alignment percentage trends, with themes like labour conditions and environmental impacts fluctuating in priority across different companies and years. At the industry level, raw material sourcing frequently emerged as a top priority, while chemical management received relatively less attention. Additionally, companies increasingly emphasized positive achievements, resulting in a noticeable positive skew in SRs, which may obscure significant challenges and raise concerns about greenwashing. This trend is particularly troubling given the industry's critical need for transparency and balanced reporting in response to significant social pressures.

This study makes significant contributions to the literature on SR analysis by introducing advanced methodological approaches and offering critical insights into corporate practices within the apparel industry. Methodologically, we advance SR analysis by employing text mining and natural language processing (NLP) techniques that address the limitations of both manual and existing automated methods. Our approach facilitates large-scale, reliable analysis of SRs, providing a comprehensive assessment of thematic alignment, prioritization, and balance, thereby establishing a new standard for SR analysis. Analytically, the study challenges prevailing assumptions by revealing that, contrary to expectations, companies within the Accord and Alliance groups prioritized environmental issues over labour concerns in the years following the Rana Plaza disaster. This unexpected thematic emphasis raises important questions about the sincerity and transparency of SR in the industry. Additionally, our research identifies a trend toward a high level of positivity in reporting across each material topic, which may serve as a red flag for potential greenwashing—whether intentional or unintentional. While we do not assert proof of greenwashing, this pattern suggests a need for cautious interpretation. Overall, this study offers a robust framework for evaluating SR practices and provides valuable insights for practitioners, regulators, and policymakers aiming to enhance the transparency and accountability of corporate sustainability efforts.

This research paper is structured as follows: the section immediately following this introduction provides a literature review of the current issue with SR, the concepts of materiality assessment and balanced reporting, as well as providing an overview of new approaches for SR analysis. Section 3

details our research methodology, providing an overview of the research context, data collection process and resulting datasets, creation of data dictionaries and the analytical tools and techniques employed for thematic extraction and sentiment analysis, aimed at assessing materiality and balance in reporting. The fourth and fifth sections present an in-depth presentation and analysis of our findings at both the company and industry levels. Finally, in the concluding section, we summarize our insights, discuss the implications of our findings for both research and practice and draw conclusions that synthesize the essence of our research endeavour.

2 Literature Review

The past two decades have seen notable shifts driven by increased attention from companies, policymakers, and stakeholders toward non-financial reporting practices. In particular, the disclosure of non-financial data has seen a substantial rise. While this trend was initially welcomed as a sign of the strategic importance companies place on engaging stakeholders through non-financial disclosures, even in the absence of legal mandates (Alliance for Corporate Transparency, 2020; KPMG international, 2014), more recent perspectives express serious concerns about their quality (Aras & Crowther, 2008; Billio et al., 2021; Boiral & Heras-Saizarbitoria, 2020; Bushee et al., 2018; Clark et al., 2019; Fabrizio & Kim, 2019; Kim & Lyon, 2015). Common among these are the transparency, relevance, and credibility of the information in SRs, which, taken together, highlight the need for better reporting practices (Cho et al., 2015; Gray, 2010; Lock & Seele, 2016). What constitutes quality sustainability reporting has been the subject of considerable debate and the subject of numerous research studies. We summarize these concerns and research contributions below, focusing on how they suggest the need to better understand two main attributes of SR quality - materiality and balance.

2.1 Concerns with Corporate Sustainability Reporting

A substantial and expanding body of research indicates that the vast quantity, variety, and intricacy of data within SRs make them difficult to evaluate, creating considerable challenges for analysis, interpretation, and comparison (Aras & Crowther, 2008; Billio et al., 2021; Boiral & Heras-Saizarbitoria, 2020; Bushee et al., 2018; Clark et al., 2019; Fabrizio & Kim, 2019; Kim & Lyon, 2015; Michelon et al., 2015). These concerns indicate that while there is an increasing volume of non-financial disclosures, the precision and utility of these reports often fall short of stakeholders' expectations.

It has been suggested that this surge in disclosure does not necessarily reflect a commitment to transparency and sustainability but rather an effort to align with institutional pressures and maintain a favourable public image (Suddaby & Panwar, 2022). Scholars observe a prevailing inclination among practitioners to disclose such information either to comply with regulations (Blasco, 2017; Ioannou & Serafeim, 2017) or to proactively respond to market pressures (Dhaliwal et al., 2011; Lyon & Maxwell, 2011). There have been serious concerns about the veracity of these reports, with the growing awareness that many companies may be engaging in "greenwashing" – misleading consumers, stakeholders, or the public by making unsubstantiated or exaggerated claims about the environmental, social, and governance (ESG) aspects of a product, service, or overall operations (Hamza & Jarboui, 2022; Hoepner et al., 2017; Kroes, 2018; Lu & Jagoda, 2023; Moodaley & Telukdarie, 2023; Santos et al., 2023; Torelli et al., 2020). Rather than transparency, some companies aim to portray a more environmentally friendly image than is justified by actual practices (Pimonenko et al., 2020).

To improve sustainability reporting practices, several actions have been undertaken and discussed in the literature. Regulatory bodies have introduced stringent guidelines like the EU's Sustainable Finance Disclosure Regulation (SFDR) to ensure transparency and accountability in sustainability claims (Hummel & Jobst, 2024; Jackson et al., 2020). Additionally, non-state frameworks like the Global Reporting Initiative (GRI) and the Sustainability Accounting Standards Board (SASB) have been developed to standardize and enhance the quality of ESG disclosures (Gamsjäger & Ray, 2024; GRI & SASB, 2021). We focus on and review the literature concerning two key points of convergence in these efforts regarding the requirements of quality sustainability reporting: the need for material and balanced disclosure. As key indicators of possible greenwashing practices, these are important determinants in scrutinizing sustainability disclosures.

2.2 Materiality in Sustainability Reporting

The concept of materiality is fundamental for transparent and accountable sustainability reporting. It refers to information that profoundly impacts a company's environmental, social, and governance (ESG) performance and its ability to create long-term value (Dhaliwal et al., 2011; Lai & Stacchezzini, 2021). In broad terms, materiality in reporting refers to the importance of the information being disclosed - the content of the sustainability report should focus on issues that are significant to both the company and its stakeholders. This includes identifying and reporting on the most critical ESG factors that could impact the company's performance and reputation (S. R. Wu et al., 2018). As such, companies are expected to engage with their stakeholders to determine what

issues are most material to them, thereby prioritizing the topics required to meet user expectations (Baumüller & Sopp, 2021; Cerbone & Maroun, 2020; Fasan & Mio, 2017; Henriques et al., 2022).

2.2.1 A recent move towards integrated materiality reporting

While the concept of materiality has its roots in financial accounting, it has evolved considerably in the context of SR. The literature on materiality in SR has seen notable policy and scholarly interest, leading to spirited debates among academics, practitioners, and standard setters about the varying definitions of materiality and the methodologies best suited to assess materiality (Fiandrino et al., 2022; Garst et al., 2022; Jørgensen et al., 2022).

Materiality in SR can generally be categorized into two main types: financial materiality and sustainability (or double) materiality. Financial materiality focuses on information that impacts a company's financial performance, while sustainability materiality considers the broader impact on society, the environment, and other external stakeholders. These different perspectives shape how companies prioritize and disclose their ESG information, influencing stakeholder perceptions and decision-making (Baumüller & Sopp, 2021).

Prominent frameworks such as the Sustainability Accounting Standards Board (SASB) and the Global Reporting Initiative (GRI) are examples of the application of different perspectives. Despite some similarities, their approaches to materiality are fundamentally different f(Adam, 2022; Pizzi et al., 2022), with GRI adopting a broad perspective that considers the interests of a wider range of stakeholders and SASB focusing on financial material sustainability information relevant to investors and (SASB, 2018; GRI, 2020). More specifically, GRI standards enable organizations to publicly disclose their most notable impacts on the economy, environment, and people, including impacts on human rights and management of these impacts (GRI, 2024). In contrast, SASB guidelines are designed to help organizations deliver industry-specific disclosures regarding sustainability-related risks and opportunities that are likely to influence the entity's cash flows, access to funding, or cost of capital across short, medium, or long-term periods. (SASB, 2024). Though SASB was initially more prominent in North America, it is now gaining significance in Europe and Nordic regions (Gamsjäger & Ray, 2021)

Some academics have highlighted the conflicting logic between GRI and SASB, emphasizing how, due to their conceptual differences in principles, orientations, and perspectives, the choice of standard influences both information disclosure and stakeholder/ investor evaluations, which affects user expectations and information needs (Jørgensen et al., 2022; Pizzi et al., 2022). Disclosing

nonfinancial information using SASB is seen as monetizing nonfinancial performance through synthetic indicators (Barby et al., 2021), whereas adopting GRI guidelines signifies a long-term vision (Adams et al., 2021).

However, the current debate on materiality is substantially influenced by international standard setters and policymakers who are working to develop new approaches that integrate both financial and sustainability materiality (Pizzi et al., 2022). These suggest that GRI and SASB should be understood as complimentary in nature and that an integrated approach, which has increasingly been adopted by a growing list of companies, is recommended for comprehensive sustainability reporting (Gamsjäger & Ray, 2024; Goswami et al., 2023; GRI & SASB, 2021; IFC, 2024; Pizzi et al., 2023). Working towards this objective, the GRI and SASB released a joint guide in 2021 that harmonized their reporting frameworks. This alignment allows companies to use both frameworks effectively, providing a holistic view of their sustainability performance (Gamsjäger & Ray, 2024; GRI & SASB, 2021). This approach ensures comprehensive sustainability reports that meet the informational needs of all stakeholders, from investors to the broader public. Companies like Nike exemplify the benefits of this dual reporting approach by integrating both GRI and SASB standards in their sustainability reports (GRI & SASB, 2021).

2.2.2 Problems with SR materiality assessment literature

Despite the availability of structured frameworks provided by GRI, IIRC, and SASB, these remain surprisingly underutilized in SR materiality assessment studies (Fiandrino & Tonelli, 2021; Michelon et al., 2022). A relatively small number of studies have attempted to align their methods with at least one of these standards (Barcellos de Paula & Gil-Lafuente, 2021; Betti et al., 2018; Calabres et al., 2019; Calabrese et al., 2015, 2016, 2017; Fasan & Mio, 2017; Gerber et al., 2024; Hajaya, 2023; Henriques et al., 2022; Khan et al., 2016; Lindgren et al., 2021b; Ortar, 2018; Padilla-Garrido et al., 2024; Slacik & Greiling, 2020; Vieira et al., 2021; Vigneau & Adams, 2023; Vouros et al., 2020; Wallbaum et al., 2011; P.-J. Wu & Huang, 2018; Zharfpeykan & Akroyd, 2023). As such, materiality assessments may not effectively capture the spectrum of issues considered material by even narrow stakeholder groups, such as the financial community. Perhaps more important is that, of these, none have attempted to integrate the use of both SASB and GRI. This likely reflects the fact that the move towards the use of harmonized frameworks has only taken place a few years ago.

Moreover, an examination of this small number of studies that do utilize structured frameworks such as GRI and SASB, reveals that the data sources and analytical approaches employed vary

substantially. These include deterministic modeling (Calabrese et al.,2016, 2017, Barcellos de Paula & Gil-Lafuente 2021, Vieira et al., 2021, Vouros et al., 2020, Wallbaum et al., 2011, Wu & Huang, 2018), surveys (Calabrese et al. 2019, 2015, Zharfpeykan & Akroyd, 2023, ortar 2018), interviews (Vigneau & Adams, 2023) and manual content analysis (Betti et al.,2018, Fasan and Mio, 2017, Gerber et al, 2024, Hajaya,2023, Henriques et al, 2022, Lindgren et al. 2021, Padilla-Garrido et al., 2024, Slacik & Greiling, 2020).

While this variety in choice of methods is not an issue in and of itself, many are narrowly focused, limited to a small group of firms, and lacking measures that robustly reflect broader stakeholder concerns. For instance, deterministic methods, such as those employed by Calabrese et al. (2016) suffer from a lack of external validity and are tailored too specifically to certain organizational contexts such as their business models, cultural operating contexts, size, and ownership focusing on SMEs in Italy, limiting their generalizability and scalability. Similarly, interview and survey methods, while insightful, often reflect inherent biases due to their design and the selection of respondents, limiting their applicability across different cultural and industrial contexts (Vigneau & Adams, 2023; Calabrese et al. 2019). Lastly, manual content analysis, although providing in-depth insights into disclosed materiality issues, is limited by its labour-intensive nature and the subjective interpretation of qualitative data (Fasan & Mio, 2017; Ngu & Amran, 2021).

As such, the examination of SR materiality remains underdeveloped, with notable weaknesses and limitations in both the use us of structured frameworks and in the use of methodologies. Overall, there are noticeable issues with reliability, scalability and reproducibility, essential features for broad applications. Findings can vary profoundly depending on the analysts' interpretations, criteria used, and methods chosen, further questioning the reliability of this approach (Krippendorff, 2019).

2.3 Balance in Sustainability Reporting

Materiality alone is however insufficient to guarantee transparency – it is also critical to understand *how* material information is communicated in sustainability reports (SRs). Balanced reporting is essential in presenting a fair and comprehensive view of a company's sustainability performance. This involves not only highlighting successes and positive outcomes but also being transparent about challenges, risks, and areas needing improvement (*GRI*, 2016). By articulating both opportunities and strengths, as well as risks and areas for improvement, balanced reporting helps build trust with stakeholders, showing that the company is committed to continuous improvement and willing to address and disclose difficult issues (Eccles et al., 2012; Vouros et al., 2020). Brammer and Pavelin

(2004) discuss how building a good reputation through transparent and balanced reporting can enhance a company's legitimacy and stakeholder trust. As such, the notion of balanced reporting is intended to counter the tendency to use SRs as promotional tools, ensuring they provide a realistic and credible account of a firm's sustainability performance (Corporate ESG Reporting, 2024; GRI, 2024; Foley, 2024; KPMG, 2023; Raghavendra et al., 2023).

Prominent frameworks like the GRI and SASB underscore the necessity for balance in reporting practices (GRI, 2024; SASB, 2021). SASB promotes a holistic consideration of both risks and opportunities, aligning sustainability issues with financial performance metrics to ensure comprehensive reporting (SASB, 2021). It emphasizes the need for the accuracy of ESG performance assessments and encourages companies to provide concise and focused reports on the most impactful sustainability issues for their business. Similarly, GRI encourages companies to disclose both risks and opportunities related to their sustainability performance, including reporting on possible risks and challenges, as well as opportunities for growth and improvement, whether or not these are financially material (GRI & SASB, 2021).

Advisory firms also increasingly point out the importance of balanced reporting. The "big four" have been particularly vocal about this aspect of SR reporting in recent years. For instance, KPMG suggests that balanced reporting should address both successes and challenges, providing a realistic view of a company's sustainability efforts (KPMG, 2023). PwC advises companies to maintain a balance by highlighting their positive contributions while being transparent about any negative impacts(Picard et al., 2023). EY echoes this sentiment, emphasizing that neglecting to report negative social and environmental impacts can threaten a company's long-term sustainability (Tomlinson & Godshall, 2023). Lastly, Deloitte advises companies to identify and report both risks and opportunities in sustainability matters (Deloitte, 2023).

Balanced sustainability reporting necessitates an examination of linguistic aspects that play a key role in communication. Overly optimistic language in SRs can undermine credibility and transparency by emphasizing strengths and opportunities while downplaying negative aspects (Du & Yu, 2021; Hamza & Jarboui, 2022; E. Kang & Lam, 2023; Mućko, 2021). This linguistic bias can mislead stakeholders about the true state of the company's sustainability performance, complicating their ability to make well-informed decisions (Hamza & Jarboui, 2022). Rigorous analysis of language and

tone is essential to ensure that sustainability reports provide a realistic and comprehensive view of both the positive and negative aspects of a company's operations.
Linguistic analysis in sustainability reporting has been examined in a few studies, revealing the significance of language use in shaping stakeholder perceptions.

For instance, Du and Yu (2021) compare the language of SRs between financial and nonfinancial firms, finding that financial firms often employ more positive language to influence perceptions. Kang and Lam (2023) explore the impact of environmental disclosures on equity underpricing, concluding that authentic and balanced language in SRs is critical for maintaining investor confidence. Hamza and Jarboui (2022) investigate the role of positive and negative language in CSR reports, emphasizing that an overly positive tone can lead to skepticism and reduced credibility among stakeholders.

Research has also highlighted the dynamics of fraudulent reporting and greenwashing, where companies manipulate information to present a favorable image (Beasley et al., 2005). This concern aligns with issues in sustainability reporting, where companies emphasize positive aspects while downplaying or omitting negative information. Clatworthy and Jones (2006) analyze the differential patterns of textual characteristics in corporate narratives, revealing how companies can manipulate language to emphasize positive outcomes and obscure negative ones. Delmas and Burbano (2011) identify the drivers of greenwashing and emphasize the importance of transparency and accountability in sustainability reporting.

For instance, Merkl-Davies and Brennan (2008) argue that narrative disclosures in SRs are often manipulated to present a more favorable view of corporate performance. They explore discretionary disclosure strategies in corporate narratives, noting how companies might incrementally release information to manage stakeholder perceptions favorably. Laufer (2003) discusses the concept of "symbolic compliance," where companies adhere to the letter of reporting standards without committing to substantive changes. This practice is further evidenced by Beasley et al. (2005), who highlight the symbolic nature of many corporate disclosures, suggesting that these reports serve more to placate stakeholders than to provide a genuine account of corporate actions.

2.4 New approaches to the analysis of sustainability reporting

Given the challenges associated with clear and transparent communication in SRs, many analysts and researchers are turning to automated text analysis approaches, such as text mining and natural language processing (NLP). These advanced methodologies are increasingly accessible and have

become valuable tools across various disciplines, including sustainability studies and the broader social sciences. Scholars in these fields have recognized the transformative possibilities of these techniques, particularly in automating the analysis of written language, which can substantially enhance research across disciplines like economics, sociology, and organizational studies (Barbier et al., 2022; DiMaggio, 2015; George et al., 2016; Leavitt et al., 2021; Tremblay et al., 2021).

The adoption of automated text analysis tools in SR analysis is driven by the need to overcome the limitations of manual analysis, such as high costs, time constraints, and the risk for human error and bias(Q. Deng et al., 2017b; Kiriu & Nozaki, 2020b; M. Li & Zhao, 2021). Automated methods enable the efficient processing of large datasets, offering a more comprehensive and objective approach to analyzing sustainability disclosures(Tremblay et al., 2015).

This research aims to apply text mining and NLP methods to extract key themes and assess the balance of reporting in SRs, contributing to the credibility and reliability of these reports. Theme extraction focuses on identifying essential topics within the reports, while sentiment analysis categorizes content based on its tone—distinguishing between positive statements that highlight strengths and opportunities and negative ones that address risks and areas for improvement. Together, these methods provide a holistic understanding of the overall sentiment conveyed in sustainability disclosures.

Despite the growing adoption of automated text analysis, there remains a lack of comprehensive studies that integrate linguistic assessment with materiality analysis. The combination of these two approaches is critical for providing a holistic view of a company's sustainability performance. While materiality assessment focuses on identifying and prioritizing important ESG issues, linguistic analysis ensures that the communication of these issues is transparent and credible. This integrated approach can enhance the overall quality and reliability of sustainability reporting, fostering greater stakeholder trust and engagement.

By leveraging these advanced analytical techniques, we aim to develop a robust framework for assessing the balance and materiality of sustainability reports, ultimately contributing to the credibility and reliability of SR practices. This endeavor aligns with the broader theoretical perspectives discussed, ensuring that SR practices not only meet but exceed stakeholder expectations, thereby promoting sustainable business practices.

2.4.1 Theme Identification/alignment analysis/BERT

Over the years, researchers have explored a variety of methodologies to assess the alignment of SRs with pre-defined non-financial reporting standards and frameworks. These methodologies range from traditional content analysis to advanced machine learning (ML) and deep learning (DL) techniques. Each approach contributes to understanding the completeness and thematic alignment of SRs, but they differ substantially in their depth, accuracy, and scalability.

- **Traditional Word Frequency and Machine Learning Techniques:**

One of the earliest methods used to analyze SRs involves word frequency analysis, which remains foundational in the field. Studies by Chang & Cheng (2013) and Aureli (2017) employed text mining techniques to identify differences in the frequency of sustainability-related words across industrial sectors and to assess changes in disclosure before and after major industrial events. These approaches utilize tools such as keyword frequency analysis and Term Frequency-Inverse Document Frequency (TF-IDF) weighting to evaluate the prevalence of certain topics to identify their alignment with reporting standards. Similarly, Tóth et al.(2022), and Wang et al. (2023) employed keyword frequencies and TF-IDF [2]weighing indicators to evaluate compliance with standards and identify sustainability-related topics associated with specific SDGs.

While foundational, these traditional methods are limited to a surface-level understanding and often fall short of capturing the complex aspects of sustainability narratives. They provide a quantitative measure of word usage but do not account for the context in which these words are used, leading to possible misinterpretations of the report's thematic focus.

Early ML approaches, such as those by Shahi et al.(2014) and Ramakrishnan et al. (2023b) employed supervised machine learning techniques like Multinomial Naïve Bayes and decision tree-based models to analyze Corporate Sustainability Reports (CSRs). These methods were particularly useful in providing structured analyses within frameworks like the Global Reporting Initiative (GRI), enabling targeted evaluations of green performance in specific sectors. However, these techniques

[2] **TF-IDF** is an advanced method that assesses the importance of terms in a document relative to a corpus. It considers both the term's frequency within the document and its rarity across the entire corpus. Terms with high TF-IDF scores are considered significant and indicative of the document's content. By identifying unique terms with high TF-IDF values, researchers can uncover specific concepts and themes that are prominent in the SRs (Raghupathi et al., 2020).

often relied on pre-labeled data and were limited by their ability to understand the context, leading to possibly superficial analyses of complex SRs.

- **Advancements with BERT and its Variants:**

As SR analysis has evolved, there has been a clear shift towards more sophisticated and context-sensitive techniques, particularly with the introduction of models like BERT (Bidirectional Encoder Representations from Transformers). Developed by Google AI and introduced by Devlin et al. (2018), BERT represents a notable leap forward in NLP. Unlike traditional models that process text in a unidirectional manner, BERT's bidirectional architecture allows it to understand the full context of a word within a sentence, capturing the complex interplay of language in SRs (Devlin et al., 2018; Stanik et al., 2021).

This advancement is particularly important for the analysis of sustainability reports, where the meaning of terms can vary greatly depending on their context. BERT's dynamic embeddings provide a nuanced understanding of textual data, making it a superior tool for dissecting complex SR content. This is a substantial improvement over static models like word2vec, which do not account for the contextual relationships between word (Devlin et al., 2018; Grootendorst, 2022).

- **Supervised and Unsupervised Learning with BERT:**

In the realm of supervised learning, studies like those by Angin et al.(2022), Li & Rockinger (2024), and Ehrhardt & Nguyen (2021) have demonstrated the effectiveness of fine-tuning BERT for specific analytical tasks. These studies harness BERT’s ability to adjust its internal parameters to the specific characteristics of the dataset, improving its accuracy in identifying and categorizing the nuanced aspects of sustainability practices.

For example, Ehrhardt & Nguyen (2021) used BERT in a supervised setting to perform Named Entity Recognition (NER), identifying entities related to coal activities in ESG reports. Angin et al.(2022) utilized the fine-tuned RoBERTa model to classify sections of SRs based on their relevance to the Sustainable Development Goals (SDGs). This approach demonstrated how BERT’s contextual understanding could enhance classification tasks. Similarly, Li & Rockinger (2024), developed an innovative SDG-related BERT model to analyze the thematic evolution of SRs in European banks. These applications show how BERT could effectively classify and track the evolution of sustainability themes, enabling faster and more accurate assessments of SRs.

While these studies effectively utilized BERT for specific tasks, they were often limited in scale, typically focusing on a relatively small number of reports (31 reports analyzed in the Ehrhardt & Nguyen (2021) study). Moreover, the necessity of labeled data restricts the application to domains where such labeled datasets are available, thereby limiting the scalability and generalizability of the analysis.

Beyond supervised learning, unsupervised techniques Latent Semantic Analysis (LSA) 3and Latent Dirichlet Allocation (LDA)4 have been used to explore the thematic alignment of SRs with frameworks like SASB. For instance, Lindgren et al. (2021b) employed LDA topic modeling to assess SRs' alignment with the Sustainability Accounting Standards Board (SASB) framework, while Bodendorf et al.(2022) used word2vec models to quantify sustainability strategies within the automotive industry. These methods automate the identification of prevalent themes and provide insights into the structure of sustainability narratives.

These unsupervised methods are scalable because they do not require labeled data, which allows for the analysis of a large number of reports. However, they often lack the contextual sensitivity that more advanced models, like BERT, can provide. For example, while LDA can identify broad topics within SRs, it tends to oversimplify the complex and nuanced language of sustainability reports, possibly leading to less precise thematic categorizations. Word2vec models, although useful for capturing word associations, do not fully account for the varying contexts in which sustainability topics are discussed, limiting the depth of the analysis.

BERT excels in uncovering and interpreting latent themes within extensive collections of unlabeled text, effectively serving as a feature extractor. This capability was exemplified in the research by

[3] **LSA** is an unsupervised method that focuses on understanding relationships between documents and the words they contain. By utilizing a technique labelled Singular Value Decomposition, LSA aims to capture hidden concepts or themes by considering the contextual usage of words. This technique allows for the identification of similar pieces of text based on the assumption that related words will appear together. LSA provides a mathematical framework to represent words and documents in a lower-dimensional space, making it easier to identify the main themes in a large text corpus, and help in tasks like document classification, information retrieval, and understanding the semantic similarity between different words (Gutierrez-Bustamante & Espinosa-Leal, 2022).

[4] **LDA** goes further in analyzing these reports. Also, an unsupervised method for document classification, it discovers natural groups of items referred to as “topics” within a corpus (i.e., collection of text, in this context a collection of SRs). LDA assumes that each document consists of a mixture of different topics, and each topic has its own set of words. The goal is to determine the distribution of topics in each document and the likelihood of specific words belonging to each topic. By employing Bayesian probability, LDA reveals the underlying topic structure and helps organize the documents into meaningful clusters (Niveditha et al., 2020; Szekely & Vom Brocke, 2017).

Kang & Kim (2022), where a pre-trained BERT model was used to generate embeddings for SRs. These embeddings were then employed in similarity searches to measure the alignment of report content with the Sustainable Development Goals (SDGs). This method allowed for a detailed and context-aware examination of sustainability reports, even in the absence of labeled training data, showcasing BERT's ability to capture the semantic richness of the text. However, a notable limitation of this study was the relatively small dataset, consisting of only 60 reports, which restricts the generalizability of their findings.

Research contribution:

Our research builds on and extends the preliminary work of Kang & Kim (2022), leveraging BERT's capabilities in an unsupervised learning context on a much larger dataset and applying it at the both company and industry level. By doing so, we aim to conduct an in-depth thematic analysis and evaluate the alignment of SRs with the SASB framework, further enriching the discourse on sustainable reporting practices. Unlike previous studies that utilized BERT primarily for classification or entity recognition, our approach focuses on extracting contextual embeddings from the text to measure the degree of alignment between the SR content and specific sustainability frameworks, such as those defined by the SASB. In this methodology, BERT is employed to generate embeddings that capture the semantic essence of the text, which are then compared to embeddings of pre-defined dictionary terms representing key sustainability topics. By calculating the similarity scores between these embeddings, we quantify the percentage of text within SRs that aligns with the predefined sustainability topics. This approach provides a precise measure of how well the reports adhere to specific sustainability standards, offering a more detailed and context-aware analysis of their content.

This method sets our research apart by focusing on the alignment of SRs with established frameworks through the use of advanced NLP techniques. By leveraging BERT's contextual embeddings, we are able to capture the nuanced language of sustainability reporting and provide a more accurate and comprehensive evaluation of how companies address key sustainability issues in their reports.

2.4.2 Sentiment Analysis Tools

Sentiment analysis has become an indispensable method for assessing the communicative balance in SRs, particularly by examining the emotional undertones in the text. This approach involves

evaluating and categorizing the emotional content within texts, assigning sentiments as positive, negative, or neutral, thus offering a detailed perspective on the underlying narrative (Cui et al., 2023; Harymawan et al., 2020; Sai et al., 2019; Smeuninx et al., 2016). Sentiment analysis has made important contributions to understanding balanced sentiments from various angles. Many studies have concentrated on analyzing public sentiment or opinion regarding corporate assertions in SRs, primarily through the analysis of social media data (Blazkova et al., 2023; Hoepner et al., 2017; Maulida & Hakim, 2022; Resnik & Koklič, 2018; Riani & Rusydiana, 2023). For instance, Blazkova et al. (2023), Hoepner et al. (2017), and Maulida & Hakim (2022) examined how public reactions on platforms like Twitter and Facebook align with corporate sustainability claims. These studies provide valuable insights into the public's perception of corporate disclosures, helping to understand the effectiveness and reception of SRs in the broader social context. These public sentiment analyses, while insightful, often miss the nuances of the SRs themselves. They focus on external reactions rather than the intrinsic balance and tone of the reports, possibly overlooking how companies internally frame their sustainability efforts an aspect that is critical to understanding quality.

Fewer studies have applied sentiment analysis directly to SRs to discern the level of balance in corporate disclosures (Harymawan et al., 2020; E. Kang & Lam, 2023; S. Kang & Kim, 2022; Lu & Jagoda, 2023). Harymawan et al. (2020) conducted sentiment analysis on 152 sustainability reports from companies in the construction sector listed on the Indonesia Stock Exchange. Their study aimed to identify trends and patterns in sustainability reporting by categorizing words into positive and negative sentiments. However, the study did not specify the exact tool used for this analysis, merely mentioning the utilization of Python software for text mining. The lack of specification on the exact sentiment analysis tool raises questions about the depth and reliability of their analysis, as different tools and algorithms can considerably impact the results. Similarly, Lu and Jaguda (2023) utilized a surface-level approach by counting positive and negative words using the Loughran and McDonald sentiment dictionary. Their methodology involved measuring the tone of sustainability reports by calculating the net tone—subtracting the number of negative words from positive words—and scaling this by the total number of words or tone words in the report. They primarily count positive and negative words without considering contextual nuances, which limits the depth of analysis.

Kang & Lam (2023) applied sentiment analysis in their investigation of the impact of environmental disclosures on IPO underpricing in Singapore. They used a basic word-counting software, likely the Linguistic Inquiry and Word Count (LIWC), to analyze the tone and authenticity of environmental disclosures. While LIWC is a widely recognized tool for textual analysis, it relies on predefined

dictionaries of words associated with specific sentiments, which can limit its ability to capture the nuanced language often found in SRs. The study demonstrated that a positive tone in environmental disclosures correlated with lower IPO underpricing, but the use of such a basic tool may oversimplify the complex language used in these reports, possibly missing out on more subtle or context-dependent sentiments.

Kang & Kim (2022) advanced the field by using BERT for sentiment analysis within sustainability reports. Their study employed BERT, specifically in a supervised learning context, using labeled data to train the model to recognize sentiment in the text. This approach marked a notable improvement over traditional methods by leveraging BERT's ability to understand context and nuance within language, allowing for a more accurate and sophisticated sentiment analysis.

Our study advances this literature by directly applying sentiment analysis to SRs and combining it with theme extraction to provide a more nuanced understanding of balance in SRs. By simultaneously analyzing both the thematic content and the sentiment associated with each theme, we provide a comprehensive perspective on how companies report their strengths, opportunities, risks, and areas for improvement. This approach yields deeper insights into the authenticity and transparency of SRs, going beyond what traditional sentiment analysis alone can achieve.

Since we did not have access to labeled data for our study, we turned to VADER (Valence Aware Dictionary and sentiment Reasoner). Among various sentiment analysis tools, VADER has emerged as a prominent tool. Introduced by Hutto and Gilbert (2014), VADER is a sentiment analysis tool that uses a lexicon and rule-based approach, particularly designed to capture sentiments. It is designed to effectively capture the nuances of sentiment expressed in the online environment, including the use of slang, emoticons, and varying intensities of emotion (Hutto & Gilbert, 2014). VADER combines a lexicon of sentiment-laden words with a set of qualitative rules to assess the sentiment of texts. Its lexicon is curated to include words and symbols frequently used in online communications, making it exceptionally adept at interpreting the sentiment of social media content (Hutto & Gilbert, 2014).

In the benchmark study by Ribeiro et al. (2016), VADER was compared with other state-of-the-art sentiment analysis methods and demonstrated commendable performance across diverse datasets. This benchmarking underscores VADER's robustness and adaptability to different contexts, reinforcing its reputation as a reliable sentiment analysis tool. The work of Thelwall et al. (2012) also supports the utility of VADER, highlighting the importance of sentiment strength detection. They

argue that tools like VADER are essential for understanding the intensity of sentiment in online discussions, which is important for accurate sentiment analysis in dynamic and expressive online environments.

Despite VADER's strengths in handling social media data, it is also highly suitable for analyzing SRs, which, although formal, often include nuanced language that VADER can effectively parse. VADER’s parsimonious approach combines both qualitative and quantitative measures, ensuring that it not only identifies the presence of sentiment but also its intensity, making it particularly useful for researchers and practitioners who require nuanced sentiment analysis (Hutto & Gilbert, 2014). The tool's ability to handle both traditional text and web-specific content like emojis and acronyms with high accuracy is a meaningful advantage.

Our study's novel contribution lies in demonstrating the application of VADER for sentiment analysis in SRs, combined with theme extraction, to offer a comprehensive assessment of the balance and transparency of these reports. By leveraging VADER's nuanced sentiment detection capabilities, we can provide a more detailed analysis of the sentiment associated with various themes in SRs.

In conclusion, VADER is a highly effective tool for sentiment analysis. Its rule-based model, tailored lexicon, and ability to capture sentiment nuances make it an ideal choice for analyzing the complex and varied text found in SRs. Our study extends the literature by showcasing VADER's utility in this context, thereby contributing to the advancement of methodologies for assessing the balance and transparency of corporate sustainability disclosures.

2.5 Conclusion of Literature Review Section

This review highlights the fragmented and often inconclusive nature of current materiality assessments, emphasizing the critical need for empirical research that enhances their integration into materiality determination processes. There is a clear necessity for further studies to develop reliable and valid methods for assessing materiality, particularly those that incorporate innovative methodologies and advanced analytical techniques. Leveraging technology for broader data integration and ensuring rigorous stakeholder engagement throughout the materiality assessment process are essential areas for future exploration.

Furthermore, the inconclusive findings in the existing literature reveal the need for empirical research focused on developing robust assessment methods for balanced reporting. Such research should

prioritize innovative methodologies that integrate advanced analytical techniques and technological advancements to enhance data integration and stakeholder participation.

Despite the recognized importance of linguistic analysis, comprehensive studies that merge linguistic assessment with materiality analysis remain scarce. This integration is vital for offering a holistic view of a company's sustainability performance. While materiality assessments are essential for identifying and prioritizing important ESG issues, linguistic analysis ensures these issues are communicated transparently and credibly. An integrated approach can considerably improve the overall quality and reliability of sustainability reporting, fostering greater trust and engagement among stakeholders.

Ultimately, focusing on both material and balanced disclosure is essential for enhancing the quality of sustainability reports. By adhering to these principles, companies can strengthen the credibility and utility of their disclosures, thereby fostering greater trust and engagement with stakeholders and contributing to more sustainable business practices.

3 Research Method

Our research method is designed to assess how a large number of companies are reporting based on themes that are common to both SASB and GRI material topics. Through theme extraction and balance analysis we systematically examine SRs in order gain valuable insights into a company's commitment to addressing material ESG issues. Theme extraction allows us to identify the emphasis companies place on each material topic, reflecting their dedication to integrating sustainability efforts into tangible business benefits and risk mitigation strategies. Moreover, we analyze the text associated with each topic to determine whether both risks and areas for improvement, as well as opportunities and strengths, are adequately addressed, as recommended by SASB and GRI's balance criteria. This comprehensive approach ensures a thorough and scalable assessment of a large number of SRs that provides the desired depth and breadth of analysis.

The first section of this research methodology section describes the research context and provides justification for these choices based on a rationale for the selection of this industry and timeperiod. A second section focuses on data collection, which is structured into two primary steps. Initially, we procured available SRs for analysis from various sources. First, we extracted 101 SRs from 15 companies affiliated with either the Accord or the Alliance, based on Jacobs and Singhal's (2017) research. These companies were chosen due to their notable garment-sourcing activities in

Bangladesh and their direct involvement or impact from the Rana Plaza disaster. Given that these companies signed the Accord and Alliance and declared their direct responsibility for the disaster, we were particularly interested in reviewing their SRs. The significance of analyzing the SRs of these companies lies in the immense scale and severity of the Rana Plaza disaster, which affected individuals and communities profoundly. Despite signing agreements like the Accord and Alliance, the literature suggests minimal post-disaster improvements in Bangladesh, emphasizing the urgent need for transparent and accountable actions. By voluntarily committing to these agreements, retailers signal their dedication to prioritizing worker safety and ethical sourcing practices. However, the effectiveness of these commitments remains debated. Therefore, we are motivated to investigate the SRs of these companies.

Additionally, our aggregate-level assessment focuses solely on the broader apparel industry. This involved analyzing 322 sustainability reports to provide a comprehensive overview of the entire apparel industry. In the second stage of data extraction, we aggregated all pertinent data on material topics provided by SASB for the apparel industry. This involved extracting comprehensive dictionaries to serve as a foundation for theme extraction. Material topics for the apparel sector were selected because the majority of the companies identified by Jacobs and Singhal (2017) belong to the apparel sector. This makes the SASB material topics for the apparel sector a reasonable benchmark.

The second section of the research methodology focuses on the tools and techniques employed for data extraction from SRs. This section elucidates the adopted research methodology and outlines the specific tools utilized to facilitate the extraction of data from Sustainability Reports.

3.1 Research Context: SR in the Apparel Industry following Rana Plaza

3.1.1 Description of research context:

The catastrophic 2013 Rana Plaza factory disaster in Bangladesh, which resulted in the loss of over 1,100 lives, is one of the most severe tragedies in the modern ready-made garment industry. This tragedy exposed severe deficiencies in safety standards and working conditions and placed the industry's supply chain under global public scrutiny, compelling stakeholders to demand greater accountability and openness from companies (Auke & Simaens, 2019).

Following this tragedy, the apparel industry faced immense pressure from civil society, governments, and the broader global community to improve sustainability practices and enhance transparency in CSR and sustainability reporting (Auke & Simaens, 2019; Jacobs & Singhal, 2017; Labowitz & Baumann-Pauly, 2014; Nast, 2023; Reinecke & Donaghey, 2015). Numerous international,

industry, civil society, and intergovernmental initiatives were created to deal with the issue. The most prominent among these, the Bangladesh Accord for Fire and Building Safety and the Alliance for Bangladesh Worker Safety, aimed at improving safety standards and ensuring accountability through greater transparency (Kayser, 2016).

Despite these decade-long efforts, improvements beyond basic safety and security measures have been limited. Investigative reports and academic literature reveal persistent governance and institutional hurdles in improving labour conditions, highlighting the difficulties in enforcing meaningful labour standards and addressing power imbalances between global buyers and local stakeholders (Ashwin et al., 2020; Rahman & Rahman, 2020). Studies by Ashwin, Kabeer, & Schüssler (2020) and Rahman & Rahman (2020) emphasize the lack of progress in implementing meaningful labour standards, while Siddiqi (2015) and Siddiqui & Uddin (2016) critique the limited scope of the Accord and Alliance, arguing that these initiatives have not adequately addressed the root causes of labour exploitation and human rights violations. Resistance from local governments and industry actors exacerbates these issues, indicating the need for stronger international cooperation and pressure (Bair et al., 2020; Barua & Ansary, 2020).

Industrial reports and surveys present a grim picture of ongoing issues in Bangladesh's garment industry. American and European demand for Bangladeshi garment production continues to rise, resulting in lower wages, precarious working conditions, and harmful environmental effects. Bangladeshi garment workers earn an estimated $25 to $75 per month, an unsustainable wage, particularly in major cities like Dhaka where most factories are located. Western corporations often manage factories through subcontractors, minimizing their direct presence and accountability, which allows them to shift blame onto subcontractors and avoid responsibility for improving working conditions (Bynum, 2021).

The existing literature indicates that governance and institutional challenges have hindered substantial improvements in labour conditions. Corporate-level critiques highlight the effectiveness of initiatives like the Accord and Alliance. The example of Loblaw illustrates the discrepancy between the company's positive SR disclosures and the legal actions against it, reflecting the limitations of SRs and their symbolic and compliance-driven nature. This hierarchical view—from broader governance challenges to specific corporate actions—reveals a notable discrepancy that requires further analysis.

3.1.2 Rationale for choice of research context

We selected the context of SR in the apparel industry following the Rana Plaza tragedy for several compelling reasons. First, the disaster led to considerable pressure on companies to increase transparency, resulting in more comprehensive disclosures and, consequently, a richer dataset for analysis. This expanded availability of data, both across the industry and over time—from 2014, the first year after the disaster, to the present—enabled us to construct a substantial dataset in terms of the number of companies involved and the longitudinal span of the data. By analyzing SRs over this period, we could evaluate how corporate reporting practices have evolved in response to the disaster.

Second, we focused on this context because, despite the intense pressure to improve the quality of SR disclosures, actual sustainability performance in the apparel industry over the past decade has been disappointing. This situation offered an objective benchmark to assess whether the reporting was genuinely material and balanced or simply a facade. Much of the previous analysis of Rana Plaza-related SRs versus actual performance has been conducted at the company level through case studies. For example, Doorey (2019) examines a negligence lawsuit against Loblaws, highlighting the shortcomings in its sustainability performance in the years following Rana Plaza. On the other hand, studies like those by Bujaki and Durocher (2020) explore how Loblaws effectively managed its image and legitimacy post-Rana Plaza through multifaceted social disclosures and SRs that were largely misleading. The comparison of these studies underscores the stark contrast between companies' claims in their SRs and their actual conduct. However, we aimed to understand whether the behavior exhibited by companies like Loblaws was representative of a broader pattern across all companies impacted by the Rana Plaza disaster and within the global garment industry as a whole. The SR data generated in the aftermath of Rana Plaza provided a unique opportunity to study corporate responses to the substantial pressures for greater SR transparency at both the corporate and industry levels.

3.2 Data collection

3.2.1 Data collection process

The data collection process for this research involved sourcing sustainability reports from multiple reliable sources. The main source we used is ResponsibilityReports.com, which is a comprehensive platform that provides instant access to sustainability reports, corporate social responsibility (CSR)

reports, and other related documents from global companies. This website boasts the most complete and up-to-date listings of responsibility reports. In addition to ResponsibilityReports.com, we conducted a thorough review of company websites to ensure comprehensive coverage of sustainability reporting within the apparel industry. This step was taken to double-check if companies had uploaded any reports that might have been missing from the ResponsibilityReports.com database.

To ensure a focused analysis within the apparel industry, specific categories were selected on ResponsibilityReports.com. The categories chosen were Apparel - Apparel Clothing, Apparel - Apparel Footwear & Accessories, and Apparel Industrial. These categories encompassed companies engaged in manufacturing apparel, clothing, footwear, accessories, and apparel-related industrial activities. By narrowing down the focus to these categories, the dataset obtained from ResponsibilityReports.com aligned with the research objectives.

Our dataset for the aggregate-level analysis includes 69 companies from the apparel industry, primarily publicly traded. These companies encompass a variety of types, including both retailers and manufacturers and are spread across multiple geographic regions, enabling a broad analysis of sustainability reporting practices. Key details for this dataset include a total of 69 companies, with 65 publicly traded and 4 privately held. Among these, 9 companies are manufacturers and 60 are retailers, emphasizing the industry's focus on consumer-facing segments. Geographically, the companies are distributed across major regions, with 13 based in Asia, 18 in Europe, 34 in North America, 3 in Oceania, and 1 in Israel. This diverse representation supports a comprehensive analysis of sustainability reporting practices across various regions and industry segments within the apparel sector.

For our company-level dataset, all companies in both the Accord and Alliance groups are publicly traded and operate as retailers. Notably, some companies primarily in the food industry also retail clothing. Geographically, all Alliance companies are based in North America, whereas most Accord companies are located in Europe, with the exceptions of PVH and Loblaw, which are based in North America.

A total of 421 reports were initially collected. However, when we categorized the reports by year, we observed that the years before 2013 had substantially fewer reports. This skewed distribution would result in unbalanced data, making it challenging to conduct meaningful comparisons in aggregate data analysis. Therefore, we decided to focus our analysis on the 322 reports available

from the year 2013 onwards at aggregate level analysis. These reports encompassed a diverse range of companies operating within the apparel industry, including both consumer goods and industrial goods sectors. The selection of these reports ensured a comprehensive and representative portrayal of sustainability practices and disclosure trends within the apparel industry from a diverse range of companies operating in different sub-sectors and geographies which enables a robust analysis of sustainability practices, trends, and performance in the apparel industry at aggregate level.

In parallel, and among the aggregate dataset, we separated 101 reports belonging to 15 companies;7 companies associated with the Accord, and 8 companies affiliated with the Alliance to perform the analysis at the company level and be able to compare two main initiatives created after the Rana Plaza disaster.

3.2.2 Overview of Datasets

A detailed overview of the two datasets is provided in the tables below:

Table 1: Overview of Aggregate-Level Analysis.

Year	**Number of Reports**	**Year**	**Number of reports**
2014	22	2018	45
2015	20	2019	58
2016	25	2020	63
2017	34	2021	55
Total number of reports at the industry level: **322**			

List of Companies analyzed at the aggregate level			
Company Name	**Public Vs Private**	**Type**	**Geography**
361 Degrees International Limited	Public	Manufacturer	Asia
Aditya Birla Fashion and Retail Ltd	Public	Manufacturer	Asia
Aritzia	Public	Retailer	North America
Boardriders	Private	Retailer	North America
Burlington	Public	Retailer	North America
Calida Group	Public	Retailer	Europe
Chico's FAS	Public	Retailer	North America
Delta Galil	Public	Retailer	Israel
Fast Retailing Co	Public	Retailer	Asia
Fourlis	Public	Retailer	Europe
Guess	Public	Retailer	North America
Kathmandu	Public	Retailer	New Zealand
Macy's	Public	Retailer	North America
PVH Corp	Public	Retailer	North America
Salvatore Ferragamo	Public	Retailer	Europe
Puma	Public	Retailer	Europe

Unifi	Public	Retailer	North America
Zalando	Public	Retailer	Europe
Accent Group	Public	Retailer	Australia
American Eagle	Public	Retailer	North America
Arvind Fashions	Public	Retailer	Asia
Bombay Rayon Fashions Limited	Public	Manufacturer	Asia
C&A	Private	Retailer	Europe
Canada Goose	Public	Retailer	North America
Children's Place	Public	Retailer	North America
Dillard	Public	Retailer	North America
Fenix	Public	Retailer	Asia
Gap	Public	Retailer	North America
H&M	Public	Retailer	Europe
Kohl's Corporation	Public	Retailer	North America
Next	Public	Retailer	Europe
Prada	Public	Retailer	Europe
Steven Madden	Public	Retailer	North America
TJX	Public	Retailer	North America
Vera Bradley	Public	Retailer	North America
Adastria	Public	Retailer	Asia
Ann Taylor Inc	Private	Retailer	North America
Ascena	Private	Retailer	North America
Bosideng International Holdings Ltd	Public	Manufacturer	Asia
Cabbeen Fashion Limited	Public	Retailer	Asia
Capri Holdings	Public	Retailer	North America
Columbia	Public	Retailer	North America
Esprit	Public	Retailer	Europe
Forward Fashion	Public	Retailer	Asia
Gerry Weber	Public	Retailer	Europe
Hallenstein Glasson	Public	Retailer	Europe
Levi	Public	Retailer	North America
Nike	Public	Retailer	North America
Ralph Lauren	Public	Retailer	North America
Super Retail Product	Public	Retailer	Australia
Tokmanni	Public	Retailer	Europe
VF Corporation	Public	Retailer	North America
Adidas	Public	Retailer	Europe
Anta Sports Products Limited	Public	Manufacturer	Asia
Bhartiya International Ltd	Public	Manufacturer	Asia
Burberry	Public	Retailer	Europe
Caleres	Public	Retailer	North America
Carter's	Public	Retailer	North America

Deckers	Public	Manufacturer	North America
European Lingerie Group AB	Public	Retailer	Europe
Fossil	Public	Retailer	North America
Gildan	Public	Manufacturer	North America
Inditex	Public	Retailer	Europe
LVMH (Louis Vuitton Moët Hennessy)	Public	Retailer	Europe
Nordstrom	Public	Retailer	North America
Ross	Public	Retailer	North America
Superior	Public	Manufacturer	North America
Toms	Public	Retailer	North America
Vipshop Holdings Limited	Public	Retailer	Asia

Table 2: Data extracted at Company level; Accord and Alliance companies.

Accord			**Alliance**		
Company Name	Number of reports	Report year	Company Name	Number of reports	Report year
Next	10	2012-2021	Canadian Tire Corporation	7	2015-2020
Loblaw	10	2012-2021	Carter's	3	2020-2022
Hennes & Mauritz (H&M)	10	2012-2021	Gap Inc.	8	2012,2014,2016, 2017-2021
Puma	10	2012,2013,2015,2016 2017-2021	Macy's	2	2018, 2020
PVH	8	2014-2021	Nordstrom	3	2018-2020
Sainsbury’s	5	2017-2021	The Children's Place	3	2019-2021
Stockmann	10	2012-2021	VF Corporation	2	2018, 2020
			Wal-Mart Stores	10	2012-2021
Total number of reports at the company level: 101					

3.3 Creation of materiality data dictionaries

3.3.1 GRI and SASB: Foundations of Sustainability Reporting

GRI and SASB serve as the cornerstone frameworks for sustainability reporting. GRI, established in 1997, provides a sector-agnostic standard aimed at informing a broad range of stakeholders, including investors, labour, civil society, and governments. In contrast, SASB, founded in 2011, offers industry-specific standards focused on financial material issues pertinent to investors and capital providers (Gamsjäger & Ray, 2024).

GRI facilitates comprehensive disclosures on an organization's impacts on the economy, environment, and society. Meanwhile, SASB hones in on the financial materiality of sustainability issues within specific industries. Studies have demonstrated that using both SASB's industry-focused standards alongside GRI's broader framework can offer a more holistic view of a company's sustainability performance (Fiandrino et al., 2022; Pizzi et al., 2021). A dual approach ensures that sustainability reports are both comprehensive and detailed, meeting the informational needs of all stakeholders (Gamsjäger & Ray, 2024; GRI & SASB, 2021; IFC, 2024).

The collabouration between GRI and SASB, highlighted by their joint guide on alignment issued in April 2021, underscores the importance of harmonizing these standards. This alignment enables companies to use both frameworks effectively, providing a holistic view of their sustainability performance (Gamsjäger & Ray, 2024; GRI & SASB, 2021). This dual approach ensures that sustainability reports are comprehensive and meet the informational needs of all stakeholders, from investors to the broader public.

In the context of this research, which is focused on the apparel industry, the industry-specific material topics identified by SASB provide a valuable starting point. We began by identifying the key material topics specified by SASB for the apparel industry. These topics are designed to address the most critical sustainability issues specific to this sector. Next, we mapped these SASB material topics to corresponding GRI General Disclosures and topic-specific standards. Additionally, we investigate a combined approach by cross-referencing the SASB material topics with GRI General Disclosures and topic-specific standards to complement the detailed guidance from SASB. This combined approach allows us to benefit from the detailed, industry-specific guidance provided by SASB while also incorporating GRI's broad, stakeholder-focused framework.

3.3.2 Overview of GRI Materiality Topics

The GRI Standards are organized into universal standards, sector standards, and topic standards (see Figure 1). The GRI 3 standards, part of the universal standards, outline material topics that assist organizations in reporting their economic, environmental, and social impacts. These material topics are essential for identifying and understanding the key issues crucial to stakeholders, as well as assessing the organization's overall impact on sustainable development.

The topic standards provide detailed reporting guidelines for specific sustainability issues, such as energy, water, emissions, and labour practices. These standards enable organizations to disclose their performance on these critical issues comprehensively. By integrating topic standards with universal and sector standards, organizations can ensure their sustainability reports are thorough and address all relevant material topics, offering a complete picture of their sustainability performance. This integration helps organizations meet the informational needs of a wide array of stakeholders, from investors to community members, enhancing the credibility and utility of their sustainability reporting.

As of the latest versions of the GRI Standards, sector standards have been developed for specific industries to provide additional context and detailed disclosures relevant to those sectors (e.g., Oil and Gas). The apparel industry is not yet among the released industries, therefore, organizations in this sector should refer to the GRI 3 universal standards and relevant topic standards for material topics (see Table 3).

Figure 3: GRI standard: Universal, Sector, and Topic standards

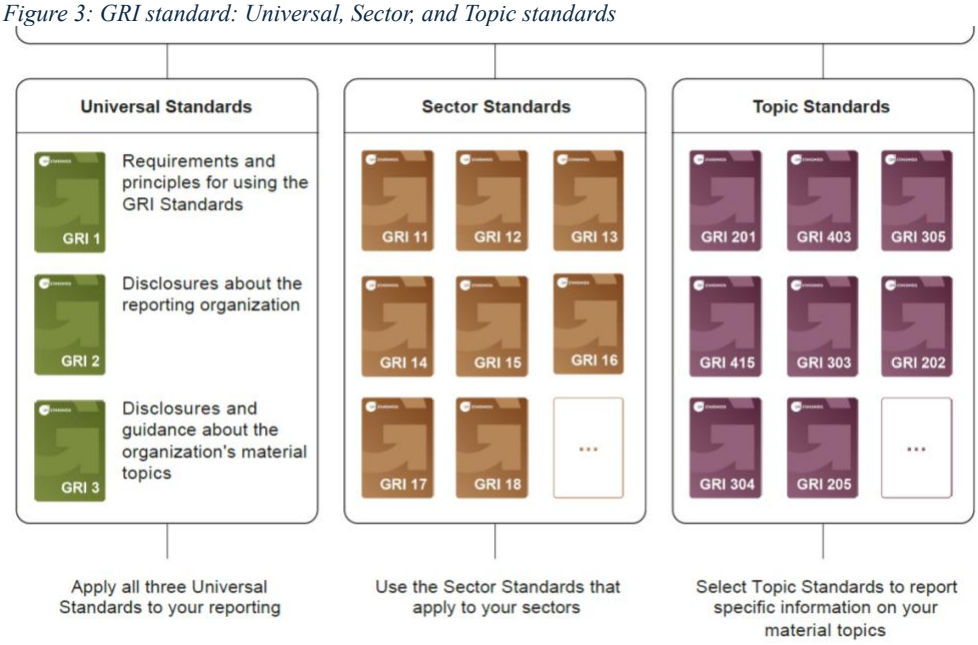


Table 3: GRI material topics and their corresponding topic standards

Category	**GRI Topic Standard**	**Description**
Economic	GRI 201: Economic Performance	Financial performance and economic contributions to local communities.
	GRI 202: Market Presence	Wages compared to local minimum wage and local hiring practices.
	GRI 203: Indirect Economic Impacts	Economic benefits are provided through investments, infrastructure, and services.
	GRI 204: Procurement Practices	Sustainable procurement policies and practices.
	GRI 205: Anti-corruption	Measures to combat corruption.
	GRI 206: Anti-competitive Behavior	Actions taken against anti-competitive behavior.
Environmental	GRI 301: Materials	Use of materials and their impact.
	GRI 302: Energy	Energy consumption and improvements in energy efficiency.
	GRI 303: Water and Effluents	Water usage, withdrawal, and discharge practices.
	GRI 304: Biodiversity	Impact on biodiversity and actions to protect habitats.
	GRI 305: Emissions	Greenhouse gas emissions and other significant air emissions.
	GRI 306: Waste	Waste generation, treatment, and disposal practices.
	GRI 307: Environmental Compliance	Compliance with environmental laws and regulations.
	GRI 308: Supplier Environmental Assessment	Environmental impact assessments of suppliers.
Social	GRI 401: Employment	Employment practices, benefits, and turnover.
	GRI 402: Labour/Management Relations	Worker consultation and negotiation processes.
	GRI 403: Occupational Health and Safety	Health and safety management and performance.
	GRI 404: Training and Education	Employee training and education.
	GRI 405: Diversity and Equal Opportunity	Diversity and equal opportunity practices.

	GRI 406: Non-discrimination	Measures to prevent discrimination.
	GRI 407: Freedom of Association and Collective Bargaining	Support for freedom of association and collective bargaining.
	GRI 408: Child Labour	Efforts to prevent child labour.
	GRI 409: Forced or Compulsory Labour	Measures to prevent forced labour.
	GRI 410: Security Practices	Training and practices for security personnel.
	GRI 411: Rights of Indigenous Peoples	Impact on indigenous communities and respect for their rights.
	GRI 412: Human Rights Assessment	Human rights risk assessments.
	GRI 413: Local Communities	Impact on local communities and engagement practices.
	GRI 414: Supplier Social Assessment	Social impact assessments of suppliers.
	GRI 415: Public Policy	Participation in public policy development and lobbying.
	GRI 416: Customer Health and Safety	Health and safety impacts of products and services.
	GRI 417: Marketing and Labeling	Marketing communications and labeling practices.
	GRI 418: Customer Privacy	Protection of customer privacy.
	GRI 419: Socioeconomic Compliance	Compliance with social and economic regulations.

3.3.3 Overview of SASB Standard

SASB has developed industry-specific standards to guide companies in disclosing financial material sustainability information. These standards are designed to meet the needs of investors and other stakeholders by providing clear and comparable data on sustainability performance. The Sustainability Accounting Standards Board (SASB) stands out for its emphasis on financial materiality, aligning sustainability issues with financial performance metrics. This focus on financially material ESG issues ensures that companies disclose information most relevant to investors seeking to assess ESG-related risks and opportunities (*SASB*, 2021). SASB's industry-specific standards provide a structured approach for companies to identify, manage, and report on material ESG issues, akin to financial accounting norms, reflecting the broader trend of recognizing ESG factors as financially material to organizational success (Busco et al., 2020).

SASB's emphasis on financially material ESG issues can help mitigate the risk of greenwashing – the deceptive practice of companies exaggerating their sustainability efforts (Moodaley & Telukdarie, 2023). Traditional sustainability reporting frameworks often lack clear guidelines on materiality, leading companies to disclose a wide range of information, not all of which may be relevant or impactful. SASB provides specific guidelines on what needs to be disclosed by emphasizing ESG topics that are both relevant and specific to each industry. Additionally, SASB insists on reporting both risks and opportunities, which compels companies to present a more balanced and transparent view of their sustainability efforts. This balanced approach to reporting

helps companies avoid making misleading sustainability claims and reduces the likelihood of greenwashing (GRI & SASB, 2021).

The SASB Standards are designed to establish and standardize disclosure criteria concerning sustainability issues deemed most critical for investor decision-making across various industries, including the Apparel, Accessories & Footwear sector (SASB website). Within this industry, the standard encompasses sustainability disclosure topics and accounting metrics related to the management of chemicals in products, environmental impacts in the supply chain, labour conditions in the supply chain, and raw materials sourcing. For the purpose of our study, we construct distinct dictionaries corresponding to each material topic specified in the SASB standard for the apparel sector. The dictionary-based approach, as elucidated by Li and Zhao (2023), provides a direct method of categorizing themes.

3.3.4 SASB and GRI materiality overlap for the Apparel industry

The integration of different sustainability reporting standards is an emerging area of interest, driven by the need for comprehensive, transparent, and comparable sustainability reports. The GRI and SASB are two leading frameworks in this domain, each offering distinct but complementary perspectives on sustainability disclosure. GRI focuses on broad stakeholder engagement and comprehensive reporting on economic, environmental, and social impacts. In contrast, SASB emphasizes financial materiality, providing industry-specific standards that cater primarily to investors and capital providers. Understanding the overlap between these two frameworks is critical for developing robust sustainability reports that address both financial and broader stakeholder concerns (ESG Seneca, 2021).

Several studies and guides have explored how companies can effectively use both GRI and SASB standards together to meet diverse stakeholder needs. For instance, the "Practical Guide to Sustainability Reporting Using GRI and SASB Standards" highlights the experiences of companies like Nike, Diageo, General Motors, and Suncor Energy, demonstrating that integrating these frameworks can offer a holistic picture of corporate performance. This approach not only enhances transparency but also ensures that critical sustainability issues relevant to both financial and broader societal stakeholders are comprehensively addressed (GRI and SASB, 2021). This combined insight underscores the complementary nature of GRI and SASB standards and provides a practical roadmap for companies aiming to integrate financial and non-financial sustainability aspects into their reporting processes. By leveraging the strengths of both frameworks, companies can better align with evolving

stakeholder expectations and regulatory requirements, thereby enhancing their overall sustainability performance and reporting credibility (Sustainability Knowledge Group, 2021).

While SASB provides industry-specific material topics for the apparel industry, these can be connected to GRI standards to create a comprehensive framework for sustainability reporting. This connection ensures that companies address both financial and broader stakeholder concerns. By mapping SASB's material topics to corresponding GRI standards, companies can enhance the robustness and relevance of their sustainability reports.

Overlapping topics:

Labour Conditions in the Supply Chain (SASB) overlaps with:

- GRI 402: Labour/Management Relations: Both address the management of labour relations.
- GRI 403: Occupational Health and Safety: Both cover aspects related to the health and safety of workers.
- GRI 407: Freedom of Association and Collective Bargaining: Both standards emphasize the importance of freedom of association and collective bargaining rights.

Environmental Impacts in the Supply Chain (SASB) overlaps with:

- GRI 301: Materials: Both focus on the environmental impacts related to material usage.
- GRI 303: Water and Effluents: Both address water usage and wastewater management.
- GRI 305: Emissions: Both cover greenhouse gas emissions and air quality.
- GRI 306: Waste: Both address waste generation and management practices.

Management of Chemicals in Products (SASB) overlaps with:

- GRI 301: Materials: Both address the impact of materials, including chemicals used in products.
- GRI 303: Water and Effluents: Both consider the impact of chemicals on water quality.
- GRI 306: Waste: Both cover the management of waste, including hazardous waste from chemicals.

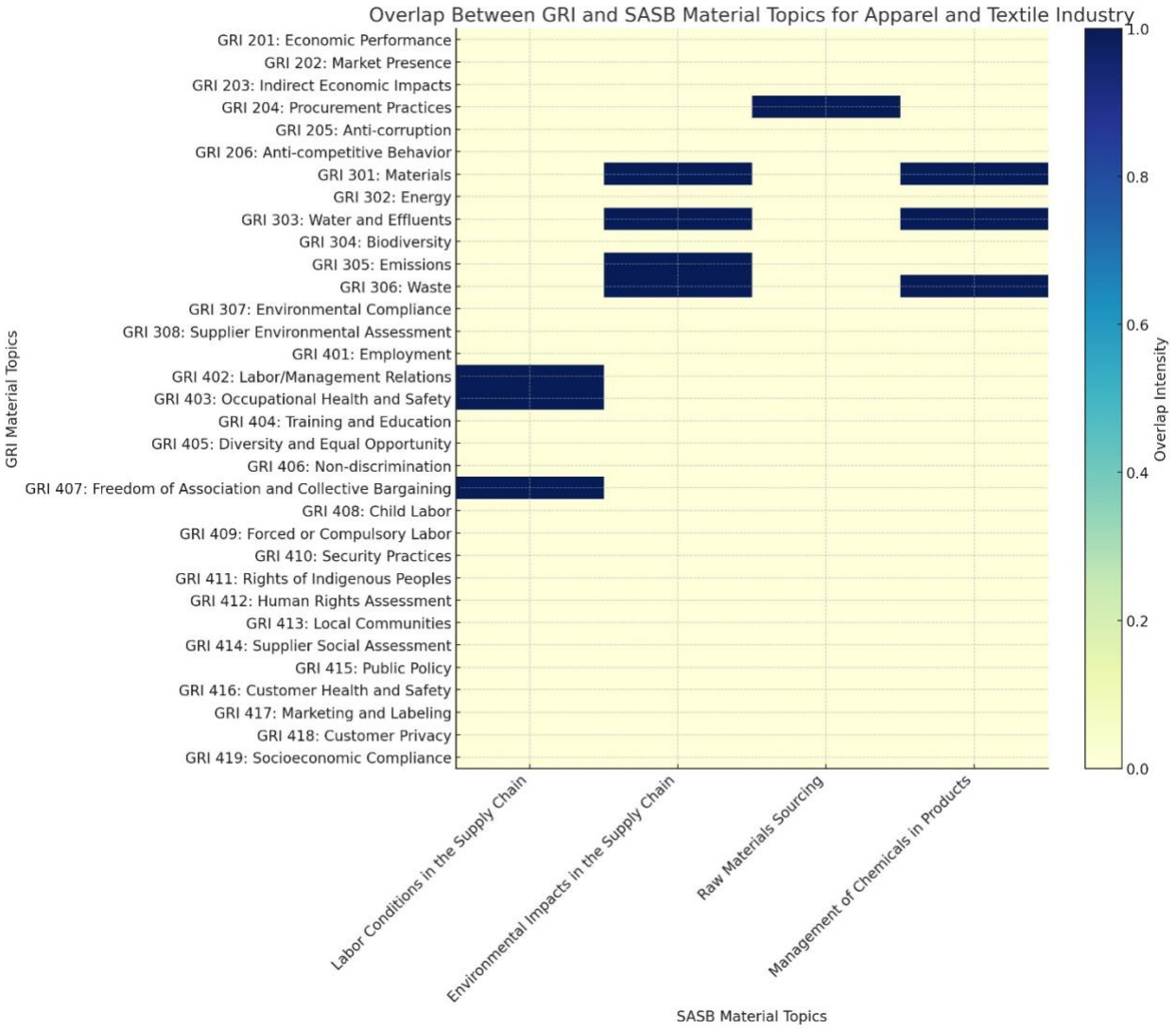
Raw Materials Sourcing (SASB) overlaps with:

- GRI 204: Procurement Practices: Both address sustainable procurement and sourcing practices.

To visualize the overlap between GRI and SASB material topics for the apparel industry, we created a heatmap (Figure 2). This heatmap clearly shows where the material topics intersect, providing a visual representation of the comprehensive coverage achieved by using both frameworks.

- **Rows** represent GRI material topics.
- **Columns** represent SASB material topics.
- **The dark color** represents the overlap

Figure 4: GRI and SASB overlap heatmap



The overlaps identified in the heatmap highlight the areas where there is alignment between what is financially material to investors (SASB) and what is broadly material to stakeholders (GRI). These overlaps represent the most critical materiality issues, capturing both financial and non-financial impacts.

The SASB topics "Labour Conditions in the Supply Chain," "Environmental Impacts in the Supply Chain," "Raw Materials Sourcing," and "Management of Chemicals in Products" cover not only financial aspects (targeted by SASB) but also broader stakeholder interests (targeted by GRI). This is

clearly demonstrated in the heatmap, which shows important overlaps with GRI topics related to labour practices, environmental impacts, and procurement processes.

These topics are comprehensive for our analysis, ensuring that companies maintain balance and transparency in their reporting. This approach aligns with both stakeholder expectations and financial performance requirements. By integrating these overlapping topics, companies can effectively communicate their ESG performance, address key concerns, and strategically manage their sustainability impacts.

Additionally, the fact that most of the SRs have employed GRI as a framework further justifies our focus on these overlapping topics. SASB, while important for financial materiality, has not been mentioned as frequently as GRI in sustainability reports, particularly before 2017. The limited mention of SASB highlights its relatively newer adoption compared to the well-established GRI. By focusing on the overlapping topics between GRI and SASB, we leverage the comprehensive stakeholder focus of GRI while integrating the financial materiality emphasized by SASB and we ensure that the most critical and relevant issues are highlighted. We are addressing key areas that impact both financial performance and stakeholder trust.

3.3.5 Data Dictionaries

For the purpose of our study, we construct distinct dictionaries corresponding to each material topic specified in the SASB standard for the apparel sector. The dictionary-based approach, as elucidated by Li and Zhao (2023), provides a direct method of categorizing themes. Building on this, to ensure the accuracy and efficacy of the dictionary, validation procedures were conducted in collabouration with a PhD student specializing in computer science and sustainability. Leveraging the BERT modeling technique alleviated the need for overly sensitive bigram or trigram word extraction or the precise capturing of synonyms. BERT's capability to capture underlying concepts facilitated the successful extraction of synonyms and related terms, even in the absence of exact matches to dictionary words. This validation process instilled confidence in the quality and reliability of the developed dictionary for subsequent analysis of sustainability reports within the apparel, accessories, and footwear industry.

Four distinct dictionaries were created based on the four overarching themes delineated in the SASB and GRI standards for the apparel industry. A summary of the four topics, along with their

corresponding dictionaries, is presented below. The SASB table provides a summary of the topics is also attached below:

Table 4: SASB material topics for the apparel and apparel industry

TOPIC	ACCOUNTING METRIC	CATEGORY	UNIT OF MEASURE	CODE
Management of Chemicals in Products	Discussion of processes to maintain compliance with restricted substances regulations	Discussion and Analysis	n/a	CG-AA-250a.1
	Discussion of processes to assess and manage risks and/or hazards associated with chemicals in products	Discussion and Analysis	n/a	CG-AA-250a.2
Environmental Impacts in the Supply Chain	Percentage of (1) Tier 1 supplier facilities and (2) supplier facilities beyond Tier 1 in compliance with wastewater discharge permits and/or contractual agreement[2]	Quantitative	Percentage (%)	CG-AA-430a.1
	Percentage of (1) Tier 1 supplier facilities and (2) supplier facilities beyond Tier 1 that have completed the Sustainable Apparel Coalition's Higg Facility Environmental Module (Higg FEM) assessment or an equivalent environmental data assessment	Quantitative	Percentage (%)	CG-AA-430a.2
Labor Conditions in the Supply Chain	Percentage of (1) Tier 1 supplier facilities and (2) supplier facilities beyond Tier 1 that have been audited to a labor code of conduct, (3) percentage of total audits conducted by a third-party auditor	Quantitative	Percentage (%)	CG-AA-430b.1
	Priority non-conformance rate and associated corrective action rate for suppliers' labor code of conduct audits[3]	Quantitative	Rate	CG-AA-430b.2
	Description of the greatest (1) labor and (2) environmental, health, and safety risks in the supply chain	Discussion and Analysis	n/a	CG-AA-430b.3
Raw Materials Sourcing	(1) List of priority raw materials; for each priority raw material: (2) environmental and/or social factor(s) most likely to threaten sourcing, (3) discussion on business risks and/or opportunities associated with environmental and/or social factors, and (4) management strategy for addressing business risks and opportunities	Discussion and Analysis	n/a	CG-AA-440a.3
	(1) Amount of priority raw materials purchased, by material, and (2) amount of each priority raw material that is certified to a third-party environmental and/or social standard, by standard	Quantitative	Metric tons (t)	CG-AA-440a.4

Source: SASB website

Management of chemicals in products. The topic summary and accounting metrics provided in the SASB standard for the management of chemicals in products within the apparel, accessories, and footwear industry highlight the increasing regulatory and stakeholder concerns regarding the use of harmful substances in consumer products. The presence of banned or regulated chemicals in finished apparel and footwear products can negatively impact human health and the environment. The topic

summary emphasizes the need for companies to effectively manage the use of chemicals of concern, develop safe alternatives, and eliminate banned substances. It also mentions the risks associated with non-compliance, including regulatory oversight, recalls, litigation, and reputational damage.

The accounting metrics outlined in the standard focus on the disclosure requirements related to the management of chemicals in products. These metrics include discussions on compliance with restricted substances regulations, verification processes, restricted substance lists, material supplier agreements, and input stream management. The metrics also cover aspects such as product testing, frequency of testing, use of third-party testing, standard operating procedures, scope of restricted substance lists, risk-based approaches to chemicals management, green chemistry principles, and certifications for verifying chemical content in products. The goal is to ensure transparency and disclosure regarding a company's efforts to manage and mitigate the risks associated with chemicals in their products. Overall, the topic summary and accounting metrics provide a comprehensive framework for companies to address the management of chemicals in products, considering both regulatory requirements and best practices in sustainability and safety.

The dictionary below includes important keywords extracted from the topic summary and accounting metrics provided in the SASB standard for the “management of chemicals in products” within the apparel, accessories, and footwear industry.

“Management of chemicals dictionary”: {restricted chemicals, Evaluation of Chemicals, Chemicals legislation, harmful substances, consumers exposure, carcinogenic substances, chemicals of concern, hazardous chemicals, Restricted substances regulations, ban the use of certain materials, labouratory testing, labouratory verification, material supplier agreements, input stream management, random sampling, standard operating procedure, restricted substance regulations, hazards associated with materials, hazards associated with substances, hazards associated with chemicals, chemicals management, toxicological characteristics, chemical ingredients, exclusion of chemicals, evaluating chemical hazard, chemical risk assessment, product design and planning, materials and chemicals procurement, finished-goods testing, product labeling, incorporate alternative chemicals, product formulation, materials substitution assessments, green chemistry principles, OEKO-TEX Standard, Eco-Passport, Eco-Certification, and toxicity information**}**

Environmental impacts in the Supply Chain. The topic of "Environmental Impacts in the Supply Chain" in the SASB standard for the Apparel, Accessories & Footwear industry focuses on the

important contribution of the industry's global supply chain to environmental externalities. These externalities include water consumption, water pollution, and air pollution resulting from the industry's energy consumption. The release of chemicals during the water-intensive processes of dyeing and tanning processes leads to water pollution, while air pollution is caused by the industry's energy consumption. These environmental impacts can harm a company's reputation and increase costs over time. The industry's reliance on manufacturers in emerging markets with weak environmental regulations worsens these problems. However, as stakeholders and consumers demand more accountability, and regulations tighten, companies are now working with suppliers to reduce their environmental impact. By improving efficiencies, cutting resource use, and controlling pollution, companies can lower costs and protect shareholder value in the long run.

The accounting metrics related to environmental impacts in the supply chain for the Apparel, Accessories & Footwear industry include two key measurements. The first metric focuses on the percentage of Tier 1 supplier facilities and supplier facilities beyond Tier 1 that comply with wastewater discharge permits and/or contractual agreements. This metric ensures that the supplier facilities meet local legal or regulatory requirements for chemical limits in wastewater discharge. Compliance is determined based on testing conducted by local officials and the company, aligning with the Sustainable Apparel Coalition's Higg Facility Environment Module.

The second metric focuses on the percentage of Tier 1 supplier facilities, as well as those beyond Tier 1, that have completed the Sustainable Apparel Coalition's Higg Facility Environmental Module assessment or a similar environmental data evaluation. This metric allows companies to evaluate the environmental performance of their suppliers. The assessment is considered equivalent if it covers all applicable categories and criteria outlined in the Higg FEM or utilizes environmental data assessments with comparable coverage.

These accounting metrics highlight the importance of managing wastewater discharge and assessing environmental performance throughout the supply chain. By monitoring compliance and engaging suppliers in environmental initiatives, companies can work towards reducing their environmental impact and ensuring sustainability in their operations.

The dictionary below includes important keywords extracted from the topic summary and accounting metrics provided in the SASB standard for the “Environmental impact in the supply chain” within the apparel, accessories, and footwear industry:

'**Environmental Impacts in the Supply Chain': {** 'environmental externalities', 'water consumption', 'water pollution', 'air pollution', 'discharge of chemicals', 'water-intensive dyeing', 'industry's energy use', 'environmental regulations', 'environmental oversight', 'environmental impact', 'resource consumption', 'limit pollution', 'resource scarcity', 'supplier facilities beyond Tier 1', 'wastewater discharge permits', 'contractual agreements', 'raw materials extraction', 'facility compliance with permits', 'Sustainable Apparel Coalition's (SAC)', 'Facility Environment Module (FEM)', 'Wastewater/Effluent', 'discharge industrial wastewater from any building', 'discharge industrial wastewater from activity', 'discharge industrial wastewater from piece of equipment', 'discharge industrial wastewater from process', 'Wastewater treatment', 'Wastewater standards', 'Discharge Hazardous Chemicals (ZDHC)', 'Programme Wastewater Guidelines', 'BSR Sustainable Water Group Water Quality Guidelines', 'The Institute of Public & Environmental Affairs(IPE)', 'wastewater guidelines', 'risks associated with discharge of water', 'Environmental constraints', 'the quality of effluent discharged to the environment', 'eliminate pollutants of concern', 'storm water discharges', 'discharge rights or permits', 'surface water', 'groundwater', 'wastewater', 'Primary treatment', 'Secondary treatment', 'utilities', 'raw materials extraction', 'Environmental Management System', 'Energy & GHG', 'Air Emissions', 'testing wastewater quality at supplier facilities', 'Tertiary treatment', 'improve water quality', 'disinfection nutrient removal' **}**

Labour Conditions in the Supply Chain. According to SASB standards regarding labour conditions in supply chains, the treatment of workers and the protection of worker rights in the Apparel, Accessories, & Footwear industry's supply chain are important and material concerns. Important elements encompass employee health and safety, fair wages, child labour, and forced labour. The industry's complex supply chain structure with multiple tiers and subcontractors makes it challenging to manage labour conditions effectively. Production often occurs in countries with limited worker protection due to low costs. This exposes companies to reputational risks, increased costs, and possible impacts on sales. Companies that emphasize robust supply chain standards, actively monitor, and engage with suppliers to address labour issues are better equipped to protect long-term shareholder value.

Accounting metrics related to labour conditions in the supply chain focus on auditing and compliance with labour code of conduct. The first metric measures the percentage of Tier 1 supplier facilities and supplier facilities beyond Tier 1 that have undergone audits to ensure compliance with labour code of conduct. Audits involve assessing various labour criteria, including working hours, minimum age

requirements, compensation practices, freedom of association, and worker treatment. The second metric evaluates the priority non-conformance rate and the corresponding corrective action rate from suppliers' labour code of conduct audits. It measures the rate of non-conformances identified and the rate at which corrective actions have been taken. These metrics help companies assess labour conditions in their supply chain and address any non-compliance issues, ensuring adherence to ethical labour practices and mitigating risks.
The dictionary below includes important keywords extracted from the topic summary and accounting metrics provided in the SASB standard for the “Labour conditions in the supply chain” within the apparel, accessories, and footwear industry:

Labour Conditions in the Supply Chain:{'treatment of workers', 'protection of worker rights', 'employee health and safety', 'fair pay', 'child labour', 'forced labour', 'part-time workers', 'regulations and enforcement protecting workers', 'labour incidents', 'labour-related work stoppages', 'strikes', 'labour concerns', 'labour code of conduct', 'working conditions', 'labour practices', 'health and safety requirements for suppliers and contractors', 'Labour criteria', 'assessment of worker hours', 'nondiscrimination', 'compensation practices', 'minimum age', 'freedom of association', 'worker involvement', 'anti-abuse policies', 'anti-harassment', 'retrenchment policies', 'occupational safety hazards', 'environmental provisions', 'human health and safety', 'Fair Labour Association (FLA)', 'risk associated with labour conditions', 'worker interview', 'Relationship with suppliers', 'development of supplier workplace code of conduct', 'worker benefit programs at supplier factories', 'harassment or abuse', 'labour conditions risks', 'incentives or mandating sanctions for labour'.**}**

Raw Materials Sourcing. According to SASB standard, the raw materials sourcing aspect of the Apparel, Accessories & Footwear industry involves the procurement of various materials, including cotton, leather, wool, rubber, and precious metals, as pivotal inputs for finished goods. Environmental and social factors such as climate change, resource scarcity, land use, and conflicts in supply chain regions pose serious challenges to the industry's ability to source materials reliably. These factors can result in risks like material shortages, supply disruptions, price fluctuations, and reputational issues. To mitigate these risks, companies can engage with suppliers, enhance supply chain visibility, adopt certification standards, and explore innovative alternative materials. Proactive management of raw materials sourcing helps companies reduce exposure to price volatility, supply disruptions, and reputational risks while opening new market opportunities.

The accounting metrics for raw materials sourcing in the Apparel, Accessories & Footwear industry are designed to gather critical information for effective management. The first metric is the disclosure of priority raw materials purchased for finished goods. It includes the identification of priority materials using the Apparel Exchange's Materials Terminology Guide and covers materials such as synthetic fibers, natural fibers, manufactured cellulosic materials, and materials derived from animals. The metric also considers vertically integrated operations and the initial amount of raw materials required for production.

The second metric highlights the amount of each priority raw material purchased that is certified to a third-party environmental and/or social standard. It encompasses standards like Apparel Exchange's Recycled Claim Standard, Global Organic Apparel Standard, Fair Trade Certified, Leather Working Group, and others. Additionally, it recognizes the scope of certified priority raw materials and the significance of materials that contribute to reliable sourcing strategies, including reclaimed and recycled materials.

These accounting metrics provide insights into the company's raw materials sourcing practices, environmental considerations, social standards, and efforts toward sustainable sourcing. By implementing strategies to address environmental and social factors, companies can mitigate risks, ensure reliable sourcing, and align with industry best practices.

The dictionary below includes important keywords extracted from the topic summary and accounting metrics provided in the SASB standard for the “Raw Materials Sourcing” within the apparel, accessories, and footwear industry:

“Raw Materials Sourcing”:**{** raw materials inputs for finished goods, materials such as cotton, leather, wool, rubber, precious minerals and metals, source materials, material shortages, priority materials, synthetic fibers, natural fibers, manufactured cellulosic materials, materials derived from animals, rayon, viscose, polyester, acrylic, spandex, nylon, foam, cashmere, mohair, flax, silk, hemp, down, Materials Portfolio, raw materials used in packaging and manufacturing, amount of materials purchased, metric tons, Apparel Exchange Recycled Claim Standard (RCS), Global Recycled Standard (GRS), Organic Content Standard (OCS), Responsible Down Standard (RDS), Responsible Wool Standard (RWS), Responsible Mohair Standard (RMS), Global Organic Apparel Standard (GOTS), Cotton Made in Africa (CmiA), Fair Trade Certified, Organic Fair Trade, Leather Working Group (LWG), Better Cotton Initiative (BCI), recycled polyester, recycled nylon, recycled lyocell.**}**

While we have made extensive efforts to encompass all relevant material topics in our dictionary, we acknowledge the possibility of missing certain terms. However, we are confident that the method detailed in the subsequent sections mitigates this risk by capturing the contextual nuances of related terms. This approach enhances the robustness of our dictionaries, allowing for effective identification of topics even in the absence of exact matches. Consequently, our method reduces the likelihood of overlooked concepts and strengthens the reliability of the analysis.

3.4 Analytical tools and techniques

Our analysis of SRs within the apparel industry following the Rana Plaza disaster follows a two-tiered approach. Initially, we conduct a focused examination of individual companies, followed by an aggregate-level exploration across the entire apparel sector. This dual-layered framework provides a comprehensive understanding of how sustainability topics are addressed at both company-specific and industry-wide levels.

At the company level, we quantified the alignment percentage with predefined material topics and analyzed the sentiment distribution of extracted sentences. This approach provides deeper insights into which topics are most frequently discussed and the balance in how each topic is addressed. By assessing both the extent of coverage and the sentiment associated with each topic, we can better understand the emphasis companies place on different aspects of sustainability and quantify the degree they represent strength and opportunity versus risks or areas for improvement.

At the aggregate level, we applied the same methodology with a modification for sentiment analysis. Specifically, we focused on extracting the top sentences that most strongly exhibited positive and negative sentiments. This allowed us to identify not only the most discussed topics but also the strongest sentiments within the broader apparel industry. By comparing the prominence and sentiment of various topics across multiple companies, we gained a comprehensive overview of industry-wide communication strategy for each material topic.

In the remainder of this section, we provide a detailed description of the analysis performed as well as the data preprocessing undertaken as a first step.

3.4.1 Data preprocessing

The collected data underwent preprocessing steps to ensure consistency and standardization. This included text cleaning, irrelevant information removal, and formatting for subsequent analysis. Key

preprocessing steps involved converting text to lowercase, tokenization, stop-word removal, and lemmatization. This phase is vital to ensure the quality and consistency of the data for subsequent analysis.

Initially, the PyPDF2 library is employed to read and parse the PDF files and extract text from each page. The extracted text is then subjected to a series of preprocessing steps to prepare it for in-depth analysis. The first step is converting text to lowercase and eliminating undesirable characters such as newline indicators. Tokenization, a critical step, involves segmenting the text into meaningful units, achieved using NLTK's tokenization functions tailored to our dataset's characteristics. Furthermore, we systematically removed stop words, employing NLTK's stop words corpus, recognizing the merit of eliminating common, non-informative words. Additionally, lemmatization, facilitated by NLTK's WordNetLemmatizer, aided in reducing words to their base form, streamlining subsequent analysis. Our approach also incorporated custom filtering mechanisms, including the removal of specific words listed in a predefined word list, filtering out numeric tokens, and retaining only alphabetic tokens of a certain length. Notably, we addressed geographical nuances by removing country names from the text data using the pycountry library. We establish empty data frames to organize both preprocessed and original sentences, ensuring clarity and transparency in our methodology. Employing a loop to iterate through each PDF file in the designated directory, we executed text extraction, preprocessing, and storage operations. Finally, we segregated preprocessed and original sentences into separate data frames, fostering a structured approach to result presentation, which is critical for comprehensibility and replicability. Through these comprehensive steps, we fortified the foundation of our analysis, poised for a robust and insightful exploration of the text data at hand.

3.4.2 Data analysis: theme identification

The data analysis in the findings section is divided into two distinct approaches: company-level analysis and industry-level analysis.

1. Company-Level Analysis:

- *Holistic Evaluation:* We evaluate how companies have addressed SASB material topics over time by quantifying their overall alignment with these topics and tracking alignment percentages across different years.
- *Depth of Disclosure:* This analysis goes a step further by quantifying the percentage alignment for each specific topic, providing a more detailed understanding of how thoroughly companies have addressed individual topics within their reports.

- *Balance Analysis:* We also analyze how balanced each company has been in addressing different SASB topics over the years.

2. Industry-Level Analysis:

- *Thematic Alignment:* This approach examines the thematic alignment of disclosures from a broader industry perspective. By analyzing how companies collectively address SASB topics, we highlight industry-wide trends.
- *Balance Analysis:* Rather than focusing on sentiment at the topic level within individual reports, we extract and analyze the most prominent positive and negative sentences for each topic at the industry level. This method provides a comprehensive understanding of the communication strategies employed across the industry, revealing how different sentiments and approaches are used to address each topic.

This dual analysis is essential to showcase the applicability of methodologies at both small and large scales and to assess how companies, particularly those linked to the Rana Plaza disaster, address material topics and balance their disclosures. This combined approach enables a robust evaluation of both individual company performance and industry-wide trends, ensuring a thorough understanding of sustainability reporting practices.

To address the "what" dimension and thematic identification, we utilized SBERT (Sentence-BERT), a variant of the BERT model optimized for sentence-level embeddings and similarity comparison. SBERT improves on BERT by fine-tuning it for semantic similarity tasks, allowing for more efficient and accurate sentence comparison. Using SBERT, we encoded sentences from sustainability reports (SRs) into embeddings and compared these embeddings to category-specific benchmarks derived from the SASB's industry-specific material topics. This approach allowed us to categorize content based on similarity scores, providing a precise measure of thematic alignment with SASB material topics such as "Management of chemicals in products," "Environmental impacts in the supply chain," "Raw materials sourcing," and "Labour conditions in the supply chain."

The SBERT model, celebrated for its capacity to capture semantic nuances, generates fixed-length vector representations (sentence embeddings) that encapsulate the contextual meaning of each sentence. By comparing these embeddings against pre-defined category embeddings created from keyword lists for each material topic, SBERT identifies sentences relevant to specific sustainability

categories. Sentences exceeding a similarity threshold of 0.6 are classified under the respective sustainability category, enabling a robust thematic analysis even if exact keywords are absent.

$$\textbf{Similarity} = \frac{\boldsymbol{A}.\boldsymbol{B}}{||\boldsymbol{A}||||\boldsymbol{B}||}$$

Here, **A** and **B** are the embeddings of two sentences. This categorization process enabled us to tabulate the number of sentences attributed to each sustainability category, providing valuable insights into the distribution of thematic content within the SRs.

3.4.3 Data analysis: sentiment analysis

Beyond thematic identification, we analyzed the sentiment balance within SRs using VADER and SBERT. Sentences were categorized as representing opportunities/strengths, neutral sentences (general information), and sentences that represent risks/ areas for improvement within each sustainability theme.

By further categorizing sentences and calculating the percentage of sentences falling into each sentiment category, our methodology offers insights into how firms portray their sustainability initiatives. By looking at the extracted sentences, this nuanced analysis enriches our understanding of corporate sustainability communication strategies.

VADER was chosen for its efficiency and accuracy in detecting direct sentiment in short texts. Although typically used for informal and social media data, it also serves well in the context of sustainability reporting due to its ability to capture nuanced sentiment. By combining VADER with BERT-based analysis, we leverage VADER's speed and straightforward sentiment categorization while enhancing its applicability to formal reports.

We employed SBERT (Sentence-BERT) to map sentences to specific sustainability categories. SBERT encodes sentences into embeddings that capture the semantic meaning, allowing for more accurate categorization and context understanding. This complements VADER's sentiment analysis by adding depth to our theme-based analysis.

While more advanced Transformer-based models like BERT or RoBERTa can capture complex language patterns and typically perform best when fine-tuned on a specific target corpus, our study lacked the labeled dataset necessary for such fine-tuning. Furthermore, implementing BERT at multiple stages would increase computational demands significantly, especially given BERT's

requirement for GPU resources and extensive processing time. Consequently, VADER on top of BERT emerged as a balanced approach, combining effectiveness with computational efficiency.

To validate VADER's performance, we applied it to Adidas' 2013 Sustainability Report (SR). Positive sentences highlighted various sustainability actions, such as water-saving technologies, virtual imaging to reduce waste, ethical labor practices, sustainable cotton sourcing, and safer chemical management. In contrast, negative sentences revealed areas for improvement, including human rights issues, labor concerns, environmental management challenges, and chemical use risks. This balance between positive and negative sentiment detection confirms VADER's effectiveness in extracting and analyzing meaningful information, accurately reflecting the company's sustainability performance and areas needing further attention.

Additionally, a manual check of Adidas' SR was conducted to verify our results, ensuring that the analysis reflects the company's sustainability claims accurately. This comprehensive validation process supports the reliability of VADER's sentiment analysis for evaluating corporate sustainability reports.

Table 5: Sentiment Analysis of Adidas' 2013 Sustainability Report

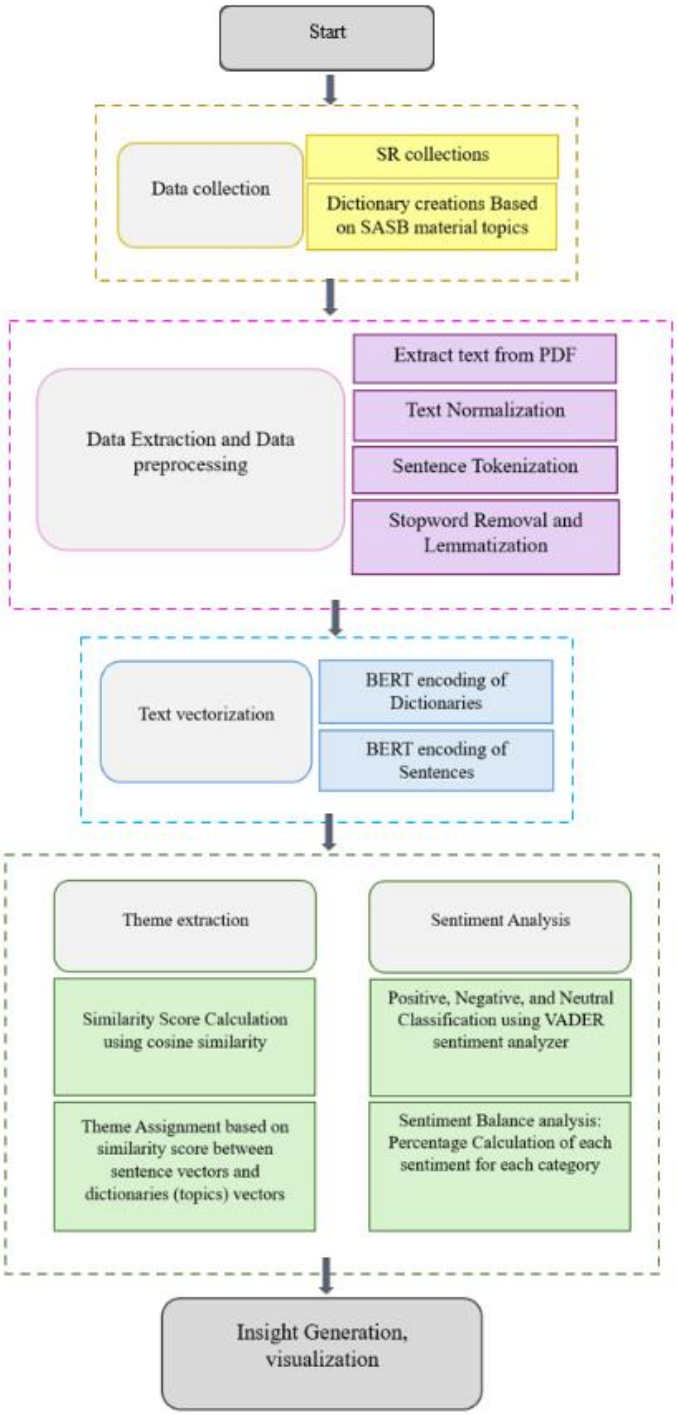
Category	Sentiment	Sentence	Analysis
Environmental Impact	Positive	"Increased use of water-saving technologies for our products, such as DryDye."	This highlights a specific action to reduce water usage, showcasing innovative sustainability practices.
	Positive	"We reduce the environmental footprint of both our own operations and our suppliers' factories."	Reflects a comprehensive commitment to minimize environmental impact across the supply chain.
	Positive	"We are reducing our environmental footprint by using virtual images to sell-in our products to our markets instead of creating physical samples."	Demonstrates innovation aimed at reducing waste and conserving resources.
	Positive	"The Framaprene® ECO heel counter material passes the strict adidas quality, fit and wear tests...This will divert 1,500 tonnes a year of polystyrene waste from landfill sites."	Illustrates a successful initiative to repurpose waste materials while maintaining product quality.
	positive	"More than three-quarters of the audited material suppliers improved their environmental performance within a year."	Highlights significant improvements in environmental performance among suppliers.
	positive	"We produced almost 1.5 million fewer physical samples, saving water, energy, and greenhouse gas emissions."	Demonstrates substantial resource savings through innovative practices.
	Negative	"While we had many successes to celebrate in 2013, we are aware that there is always more that we can do."	Acknowledges the need for continual progress in sustainability without mixing in positive statements.
	Negative	"For a number of factories, the 2013 E-KPI scores had to be	Indicates issues with data accuracy and consistency that

		readjusted due to data inconsistencies reported into our environmental data system in previous years."	need to be addressed. (areas for improvement)
Social/Labour Aspects	Positive	"We positively influence the lives of our employees, factory workers and people living in the communities where we have a business presence."	Highlights a broad commitment to improving working conditions and community well-being.
	Positive	"We have created a new Health Management department...to create a sustainable, healthy community."	Demonstrates a proactive approach to fostering a healthier work environment.
	Positive	"Systematically promote motivation and accountability through talent and performance management."	
	Positive	"Raise the overall understanding and awareness of health and safety issues specific to the manufacturing process within the supply chain."	
	positive	"25% of our strategic suppliers to be in a self governance compliance model (4C or better)."	
	positive	"54% of our direct suppliers met 3C rating (or better)."	
	Negative	"We need to know whenever we are having a negative impact on someone's human rights whether it relates to our employees, our customers, the workers in our suppliers' factories or the communities in which we operate."	Highlights a potential risk area concerning human rights impacts, focusing on vigilance and responsiveness.
	Negative	"Other findings were less positive. There were instances of under-payment of overtime, workers not being well informed of the current legal minimum wage, the piece-rate system being the overly dominant method of calculating wages, and the lack of a direct link between the wages paid and the factory's economic performance."	Provides a detailed description of specific labour issues, ensuring no positive spin on these negative findings.
	Negative	"Despit our support through capacity building, we faced challenges driving improvement in newly accepted supplier factories, and in factories where the adidas Group has limited business leverage (i.e. small business volume)."	
Raw Material Sourcing	Positive	"In 2013, the adidas Group sourced more than 23% of all our cotton as Better Cotton, clearly exceeding our milestone of 15%."	Demonstrates strong progress towards sustainability goals, supporting better agricultural practices.
	Positive	"By 2018 we have committed to source 100% of cotton in our products as sustainable cotton."	Indicates a strong dedication to sustainable sourcing with a clear and ambitious future goal.
	Positive	Through our commitment to the Better Cotton Fast Track Program (BCTFP), we are working with other leading brands and funders to help make	

		sustainable cotton production mainstream.	
	Positive	The BCFTP funds training for cotton farmers to help them adopt more sustainable practices, which encourages wider production of Better Cotton.	
	Positive	In 2013, we sourced 96% of our non-European leather volume from tanneries that achieved a LWG Silver or Gold rating. Compared to 2012, we increased our sourcing from Goldcertified tanneries by 8% to 87%, and reaching our 2015 target	
	Negative	"We want them to tackle the environmental impacts from the manufacturing process and be smart in their use of materials – while still meeting our quality standards."	Points out ongoing challenges in ensuring suppliers manage their environmental impact effectively.
Management of Chemicals	Positive	"We have developed a better way to make our products water-repellent, one that avoids using a particular group of chemicals that have been linked with potential risks to health and the environment."	Emphasizes a proactive initiative to replace harmful chemicals with safer alternatives.
	Positive	"We replaced hazardous glues with water-based chemicals in our athletic shoes and have led in the development of technical training to help build knowledge of health and safety in countries such as China and India."	Highlights the commitment to reducing health and environmental risks associated with chemical use.
	Negative	"The adidas Group was the first global brand to introduce a total ban on key hazardous chemicals in the workplace."	Implies past risks associated with chemical management that required banning hazardous chemicals.

The various steps of our methodology are illustrated in Figure 5, which presents a flowchart detailing the complete methodological process.

Figure 5: Research Framework



4 Findings - Company level

The thematic identification of SASB material topics is explored at the company level. Using the text mining methodology described above, we identified and categorized sentences relevant to these relevant topics across different years. By categorizing and analyzing sentences across multiple years, we observe variations in thematic focus and sentiment balance. Thematic alignment and sentiment distribution visualizations reveal the emphasis companies place on different sustainability aspects.

4.1 Theme extraction-company level

4.1.1 Holistic Evaluation: Comparative Analysis of SASB Topic Integration in SRs

In our detailed analysis (Table 6), to quantitatively measure the extent to which companies addressed SASB material topics over time, we employed metrics such as the Total Number of Sentences Analyzed (TNSA) and the Total Number of Aligned Sentences (TNAS). TNSA provides a count of sentences in each report per year, whereas TNAS identifies sentences that align with SASB's material topics. By calculating the Alignment Percentage (AP), which is the ratio of TNAS to TNSA, we assessed the extent to which companies addressed these topics over time. This quantitative measure allowed us to track changes and trends in sustainability reporting practices, providing a framework for our analysis.

A line chart visually represents AP trends (figure 6), showing each company's year-on-year alignment. This chart depicts the fluctuation in alignment percentages across years for various companies. Each line on the chart represents a different company, with the y-axis indicating the alignment percentage and the x-axis representing the years under review. This visualization clearly explains how companies' alignment with SASB material topics evolved over the analyzed period.

These findings illuminate the extent to which the reports include information on four SASB material topics. Most companies show an increasing trend and a steady improvement in aligning their reports with key sustainability topics indicating a growing commitment. This general upward trend suggests these topics has become a more integral part of corporate strategy and reporting practices.

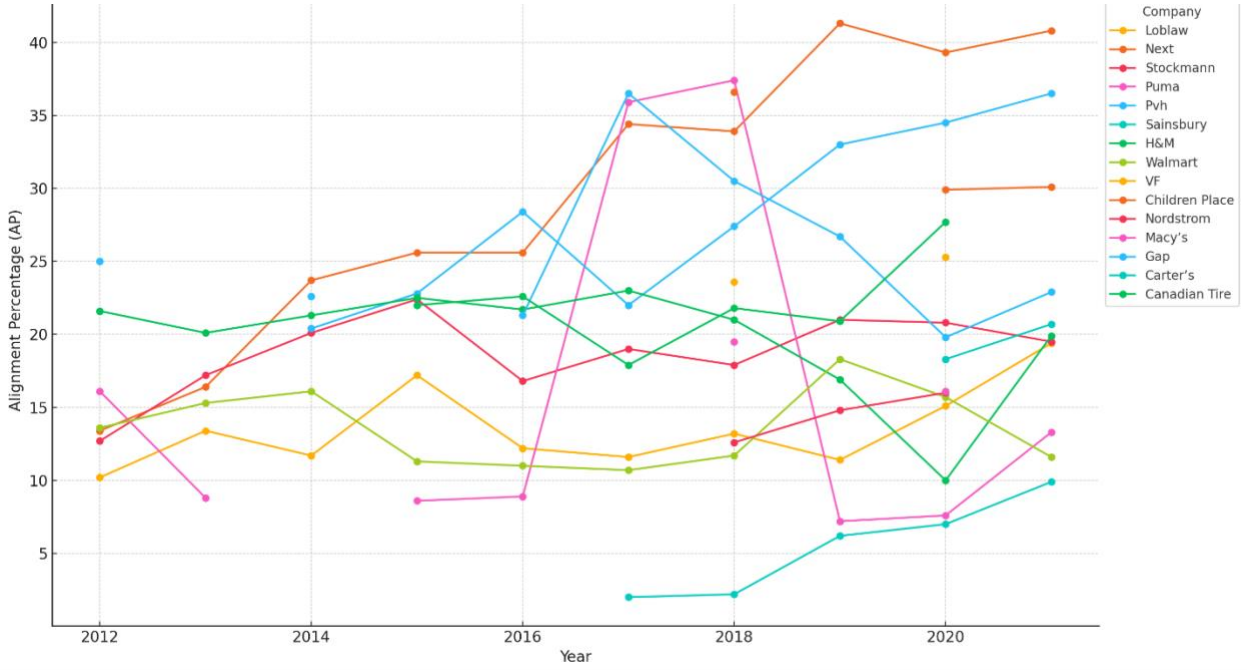
Some companies exhibit fluctuations in their alignment percentages, which could suggest changes in strategic priorities, variations in reporting standards, or responses to external pressures and stakeholder expectations. While the general trend across most companies shows an enhanced focus on these sustainability topics over time, the variations and differences in alignment percentages

highlight that while progress is being made, there is still variability in how different companies prioritize and report on these sustainability topics.

Table 6: Alignment Percentage (AP) with SASB Material Topics Over Years

Year	SR Data	Loblaw	Next	Stockmann	Puma	Pvh	Sainsbury	H&M	Walmart	VF	Children Place	Nordstrom	Macy's	Gap	Carter's	Canadian Tier
2012	TNSA	754	1004	620	2191	-	-	1086	1926	-	-	-	-	1683		-
	TNAS	77	135	79	353	-	-	235	263	-	-	-	-	421		-
	PA	10.2%	13.4%	12.7%	16.1%	-	-	21.6%	13.6%	-	-	-	-	25%		-
2013	TNSA	443	898	814	1988	-	-	1115	1933	-	-	-	-	-		-
	TNAS	58	147	140	174	-	-	224	295	-	-	-	-	-		-
	PA	13.4%	16.4%	17.2%	8.8%	-	-	20.1%	15.3%	-	-	-	-	-		-
2014	TNSA	477	689	1017	-	653	-	1538	2063	-	-	-	-	1282		-
	TNAS	56	163	205	-	133	-	327	333	-	-	-	-	290		-
	PA	11.7%	23.7%	20.1%	-	20.4%	-	21.3%	16.1%	-	-	-	-	22.6%		-
2015	TNSA	395	648	857	2014	545	-	1779	1511	-	-	-	-	-		313
	TNAS	68	166	192	174	124	-	401	171	-	-	-	-	-		69
	PA	17.2%	25.6%	22.4%	8.6%	22.8%	-	22.5%	11.3%	-	-	-	-	-		22%
2016	TNSA	378	648	1083	2181	605	-	1906	1822	-	-	-	-	1162		318
	TNAS	46	166	202	195	172	-	413	200	-	-	-	-	248		72
	PA	12.2%	25.6%	16.8%	8.9%	28.4	-	21.7%	11%		-	-	-	21.3%		22.6%
2017	TNSA	395	282	898	340	567	3266	1366	2049	-	-	-	-	622	-	485
	TNAS	46	97	171	122	125	65	314	220	-	-	-	-	227	-	87
	PA	11.6%	34.4%	19.0%	35.9%	22.0%	2.0%	23.0%	10.7%	-	-	-	-	36.5%	-	17.9%
2018	TNSA	228	319	922	334	683	3196	1730	2393	1098	153	286	508	738	-	225
	TNAS	30	108	165	125	187	69	363	280	259	56	36	99	225	-	49
	PA	13.2%	33.9%	17.9%	37.4%	27.4%	2.2%	21.0%	11.7%	23.6%	36.6%	12.6%	19.5%	30.5%	-	21.8%
2019	TNSA	324	373	789	3050	627	470	1234	830	-	-	318	-	944	-	273
	TNAS	37	154	166	220	207	29	209	152	-	-	47	-	252	-	57
	PA	11.4%	41.3%	21.0%	7.2%	33.0%	6.2%	16.9%	18.3%	-	-	14.8%	-	26.7%	-	20.9%
2020	TNSA	773	443	800	2807	919	330	2565	1140	1150	900	413	361	647	540	267
	TNAS	117	174	166	212	317	23	257	179	291	269	66	58	128	99	74
	PA	15.1%	39.3%	20.8%	7.6%	34.5%	7.0%	10.0%	15.7%	25.3%	29.9%	16.0%	16.1%	19.8%	18.3%	27.7%
2021	TNSA	777	534	834	3254	747	718	1379	371	-	948	-	-	795	644	-
	TNAS	151	218	163	432	273	71	275	43	-	285	-	-	182	133	-
	PA	19.4%	40.8%	19.5%	13.3%	36.5%	9.9%	19.9%	11.6%	-	30.1%	-	-	22.9%	20.7%	-
2022	TNSA	-	-	-	-	-	-	-	-	-	-	-	-	570	-	-
	TNAS	-	-	-	-	-	-	-	-	-	-	-	-	167	-	-
	PA	-	-	-	-	-	-	-	-	-	-	-	-	29.3%	-	-

Figure 6: Yearly Alignment Trend with SASB Material Topics by Company for Alliance Group



4.1.2 Depth of Disclosure: Assessing Topic-Specific Engagement in Accord and Alliance Sustainability Reports

In this section, we aim to assess companies based on the depth of reporting on each of the four material topics outlined by SASB, using stacked bar charts to visually delineate the distribution of reporting depth across these topics for each year (see Figure 7). Each series of bars represents a specific year, each colour corresponds to one of the material topics. It is important to note that the percentages depicted in the stacked bar charts do not reflect the distribution based on the total number of sentences in the reports. Instead, they signify the distribution among the subset of aligned sentences investigated in the preceding section or the Total Number of Aligned Sentences (TNAS).

We further provided a line plot with trend lines for each topic over the years, showing how the alignment percentages have changed for each topic (see Figure 8). This visualization will provide insights into the trends across the different companies.

This visualization facilitates a comparative analysis of reporting depth across the material topics and enables the identification of trends or shifts in companies' reporting priorities over the analyzed period. By systematically assessing the distribution of reporting across these critical sustainability themes, our study provides valuable insights into the transparency and accountability of apparel companies regarding their environmental and social performance.

Figure 7: Alignment Percentage By year and topic for each company

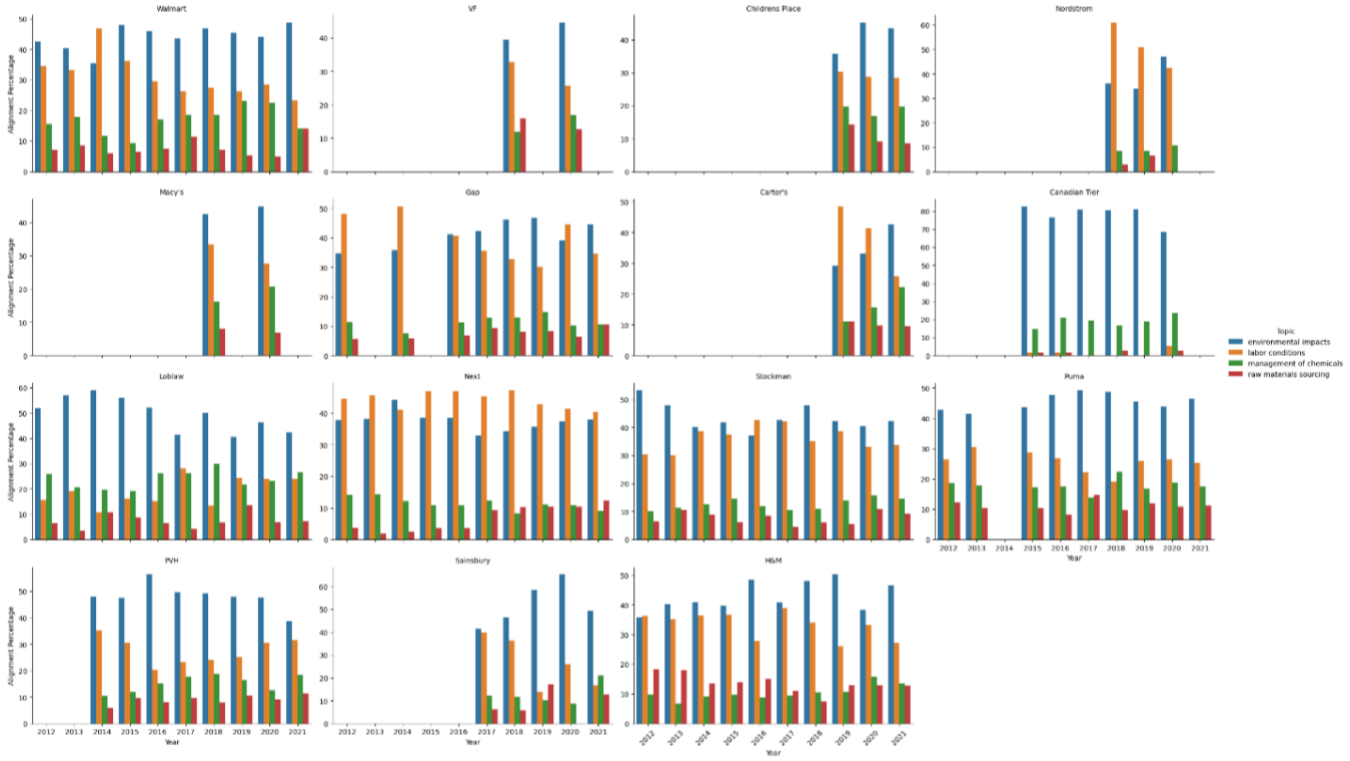
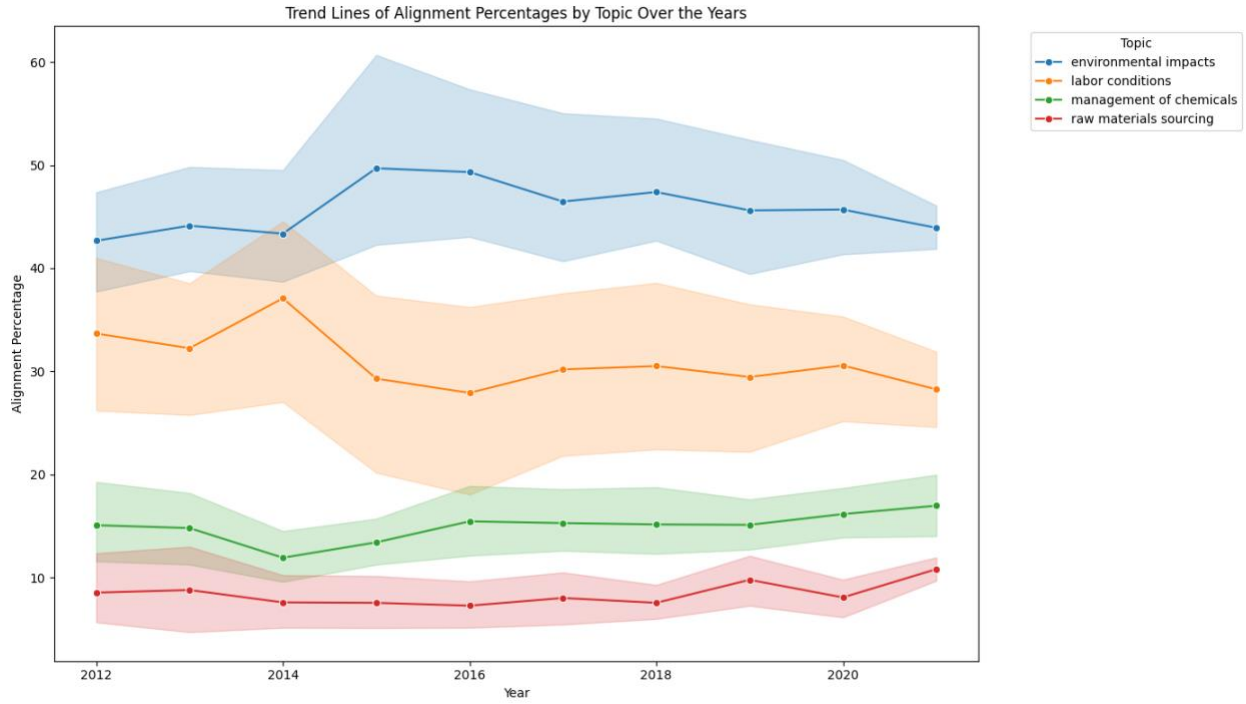


Figure 8: Trend lines of AP by topic over the years



The general trend across most companies shows different companies exhibit varying levels of alignment across the four topics. Some companies have a consistent reporting pattern, while others have notable fluctuations.

Generally, environmental impacts have higher alignment percentages compared to other topics across most companies. Companies like Walmart and Loblaw show high and relatively stable alignment in environmental impacts over the years. Labour conditions show moderate to high alignment percentages, with some companies like Walmart and Gap having noticeable fluctuations. Some companies report consistently high alignment (e.g., Gap) while others have considerable gaps (e.g., Canadian Tier).

Management of chemical topics has lower alignment percentages compared to environmental impacts and labour conditions.

Companies like Loblaw and Puma have higher alignment percentages in certain years, indicating a periodic focus on chemical management. Also, the Raw material sourcing topic generally has the lowest alignment percentages. H&M has shown a noticeable increase in raw materials sourcing alignment in recent years, indicating a growing focus on this area.

Overall, environmental impacts and labour conditions appear to be more consistently reported across companies, indicating these are well-established focus areas in sustainability reporting. Environmental impacts consistently have the highest alignment percentages across the years, showing a stable but slightly declining trend. Labour conditions also show a stable trend but with more variability compared to environmental impacts. This could reflect changes in internal policies, external pressures, or varying stakeholder priorities. Management of chemicals and raw materials sourcing shows more notable variability, indicating inconsistent reporting or varying focus among companies. This variability could be due to changes in strategic priorities, reporting standards, or external influences. Moreover, there has been a noticeable increase in alignment percentages for raw materials sourcing in recent years, suggesting a growing industry focus on sustainable sourcing practices.

The detailed trend analysis provides valuable insights into the evolving focus of companies on key sustainability topics. The consistent improvement and high alignment percentages in environmental impacts and labour conditions suggest that these areas are integral to corporate sustainability strategies.

4.2 Balance Reporting

In the preceding section, we examined the extent to which sustainability reports align to SASB standards, conducting both a comprehensive evaluation and a detailed analysis by topic. We now turn our scrutiny toward discerning whether these reports present a balanced portrayal of corporate sustainability efforts or if they are mere veneers masking less favorable realities. This examination is pivotal to understanding the authenticity and thoroughness of the companies' sustainability narratives.

The concept of balanced reporting is central to our analysis. It involves assessing whether companies present both the positive developments and the challenges in their sustainability practices with equal emphasis. In order to quantify the sentiments associated with various sustainability topics, we have developed two specialized heat maps—one for the Accord group and another for the Alliance group. These visual tools, crafted from comprehensive thematic and sentiment data, not only indicate the degree of alignment with SASB topics over time but also illustrate the balance of positive and negative sentiments within these alignments.

As articulated in the previous section, the "Category Percentage" (AP) reflects the extent to which each report addresses a given sustainability topic. Delving deeper into this category data, we further categorize sentiments into positive, negative, and neutral. Here, the Positive Percentage reflects the part of reporting on a specific topic that casts the company's actions or outcomes in a good light. Conversely, the Negative Percentage underscores the part that brings attention to the company's challenges or adverse outcomes within that sustainability domain.

We introduce a "Balance Score" metric to evaluate the equilibrium between the positive and negative disclosures, factoring in their proportion relative to the total discussion of each topic. This metric aids stakeholders in assessing whether a company's reporting might be disproportionately favorable or whether it candidly recognizes its shortcomings. Such a balanced method of scrutinizing corporate disclosures promotes a fuller understanding of a company's sustainability conduct and openness.

- **Calculation of Balance Score:**

The annual Balance Score for each company and topic is computed using this formula:

$$\text{Normalized Balance Score} = \frac{\text{Positive \%} - \text{Negative \%}}{\textbf{Category Percentage}\%}$$

The score is normalized by dividing the net positive (Positive Percentage minus Negative Percentage) by the Category Percentage. This normalization is critical; it weighs the importance of the topic in the report and ensures that more frequently discussed topics proportionately affect the Balance Score. By normalizing the balance score in relation to the category's presence in the report, we ensure that the score accurately reflects the company's reporting behavior.

This score spans between -1 and +1, where:

- A score of 0 implies equal representation of positives and negatives, indicating balanced reporting.
- A score nearing +1 signals a possibly overly positive portrayal, which might suggest underreporting of negatives or possible greenwashing.
- A score nearing -1 reflects a primarily negative depiction, which can either highlight areas for improvement or point to a company's transparency about its challenges.

Our heatmap displays these scores, allowing a quick visual interpretation of the balance in companies' reporting across various sustainability topics over time:

- Green shades represent scores closer to 1, indicating a more positive balance.
- Red shades indicate scores closer to -1, suggesting a negative balance.
- Yellow or colours close to the center of the spectrum denote scores near 0, suggesting a balanced reporting perspective.

The heatmap's intuitive colour-coding facilitates a swift visual assessment, enabling stakeholders to identify exemplary practices (green), areas needing focus (red), and balanced disclosures (yellow).

Overall, this approach to calculating and visualizing balance scores provides a nuanced picture of corporate sustainability reporting, facilitating an informed analysis of companies' practices and transparency.

4.2.1 Balance Analysis – Company Level

The analysis of balance scores, as displayed in Figure 9, reveals important insights into the sustainability reporting practices of analyzed companies. Very few instances show balance scores around 0, which would indicate an even discussion of both positive and negative aspects of what we earlier referred to as opportunities and risks.

Categories like 'labour conditions' and ‘Environmental impacts' often have higher balance scores across multiple companies and years, indicating a trend toward emphasizing positive actions and achievements. For instance, **Next** shows a consistent trend of higher scores for 'labour conditions' and ‘Environmental impacts.' However, **Management of chemicals** and **raw materials sourcing** tend to have lower and more variable scores across companies, indicating these are areas where challenges are more frequently discussed compared to 'labour conditions' and ‘Environmental impacts.'

Scores below 0, indicating a focus on challenges and issues, are less common but do appear in certain companies and years. **Canadian Tire** and **Carter's** show negative scores in some years and categories, indicating they are discussing more about challenges and issues in those areas.

There is a noticeable variation across the years for some companies. For example, **Puma** shows fluctuating balance scores, which could suggest changes in their reporting strategy or actual performance variations.

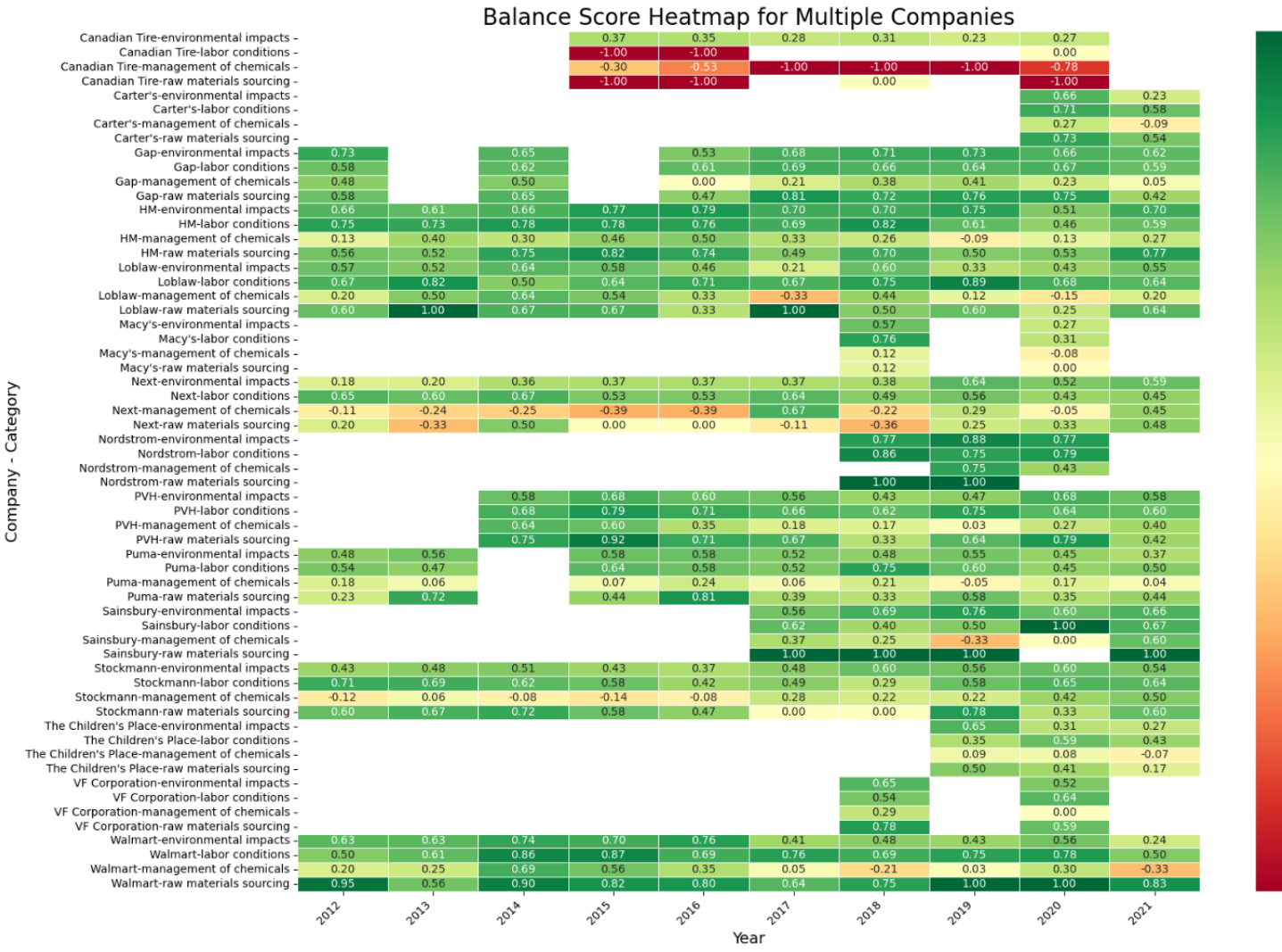
There is a noticeable increase in balance scores over time for many companies, indicating that reports are becoming more positive. The early years (2012-2015) show more variability in balance scores, with more instances of negative and neutral scores. This could suggest that companies were more candid about challenges in their early reports. 'Environmental impacts' and 'labour conditions' are categories that have shown notable changes over time. These categories often move from more balanced or negative scores in early years to highly positive scores in recent years.

This trend can be observed in companies like **Next**, **Sainsbury**, **PVH**, and **H&M**. For example, **Next** shows an increase in balance scores from negative or neutral in 2012 to highly positive by 2021. Also, **Next** and **Puma** have several negative scores in the early years.

The improvement in balance scores over time could also be attributed to better reporting practices, increased awareness of sustainability issues, and possibly efforts to present a more favourable public image.

Overall, the data indicates that over time, from 2012 to 2021, companies are increasingly reporting more positive aspects of their sustainability efforts. This trend could be indicative of an improvement in sustainability practices, better-reporting standards, or a strategic shift towards highlighting positive outcomes. However, the trend also raises questions about the likelihood of greenwashing, as many companies are moving towards presenting a highly positive narrative. The category of 'management of chemicals' remains a challenge, as it often shows lower or more variable scores, suggesting ongoing issues in this area.

Figure 9: Balance score heatmap for accord and alliance companies



4.3 Overview- Company level

Regardless of whether companies report extensively or minimally on material topics, the percentage of positivity in reporting remains consistently high across the board. While this uniform positivity might be anticipated, it raises serious concerns about the authenticity and balance of these reports. The positive skew in reporting can be seen as an effort to attract investors by portraying the company in the most favourable light, which may not be entirely accurate. This trend raises a red flag, suggesting a likely emphasis on highlighting achievements while underreporting negative outcomes, thereby raising concerns about greenwashing.

Although such positivity in reports is expected, it is vital to highlight this red flag as the literature indicates that the situation in major sourcing countries like Bangladesh remains dire more than a decade after the Rana Plaza disaster, with these companies being the largest retailers sourcing from Bangladesh. Despite the apparent positivity in reporting on these topics, the on-ground realities do not reflect substantial progress. This discrepancy suggests that the positive statements in sustainability reports may be more symbolic than substantive, necessitating further analysis by future researchers.

This observation prompts further investigation at the aggregate level, encompassing a broader range of companies. By examining the alignment of priority topics on this broader scale, we aim to determine whether the observed trends hold consistently across the industry. Anticipating a generally positive tone, we will also randomly extract the top positive and negative sentences to analyze the communicative strategies employed. This approach will provide deeper insights into how companies present their sustainability efforts and the balance between highlighting achievements and acknowledging challenges.

This comparison underscores the necessity for stakeholders to critically evaluate SRs, ensuring that high balance scores are not misleading and that they reflect a comprehensive picture of a company's sustainability efforts. Stakeholders should be particularly vigilant when scores suggest a positive skew, as it could mask areas needing improvement or signal an attempt to project an overly favorable image that may not align with the full reality of the company's environmental and social impacts.

5 Findings – Industry level

5.1 Materiality at the industry level

By expanding our analysis to include over 322 companies in the apparel industry, we have gained a comprehensive understanding of their alignment with SASB material topics. This broader perspective provides valuable insights beyond the Accord and Alliance groups, encompassing a diverse range of companies with varying strategic priorities.

At the aggregate level, we applied the same methods used at the company level, focusing specifically on companies outside the Accord and Alliance groups. Although Accord and Alliance are predominantly known for being major buyers of garment suppliers in Bangladesh, they span various

sectors, including food and home hardware. This sectoral diversity means their SRs do not fully represent the entire apparel industry regarding alignment with SASB material topics. In this analysis, we concentrated exclusively on the apparel industry, using the same rigorous methodology.

By examining the SRs of large number of SRs, we gained a comprehensive understanding of how they align with the SASB material topics. This broader analysis provides a more complete picture of the industry. The aggregate data reveals key trends and shifts in sustainability focus over time, offering a nuanced understanding of industry-wide practices.

The analysis of alignment percentages of aggregate data (table 7) we can conclude that, the alignment percentage for the management of chemicals topic has consistently been the lowest, ranging from 2.79% to 3.46%. This suggests that while companies recognize the importance of managing chemicals, it remains a lower priority compared to other sustainability topics. The consistent yet low percentage indicates incremental efforts in this area, overshadowed by more pressing issues such as labour conditions and environmental impacts.

Environmental Impacts in Supply Chains have seen steady and strong attention, with percentages ranging from 4.07% to 4.96%. The steady attention to environmental impacts reflects the industry's ongoing commitment to sustainability and environmental accountability. This consistency aligns with broader global trends where environmental issues are increasingly scrutinized by regulators, consumers, and stakeholders.

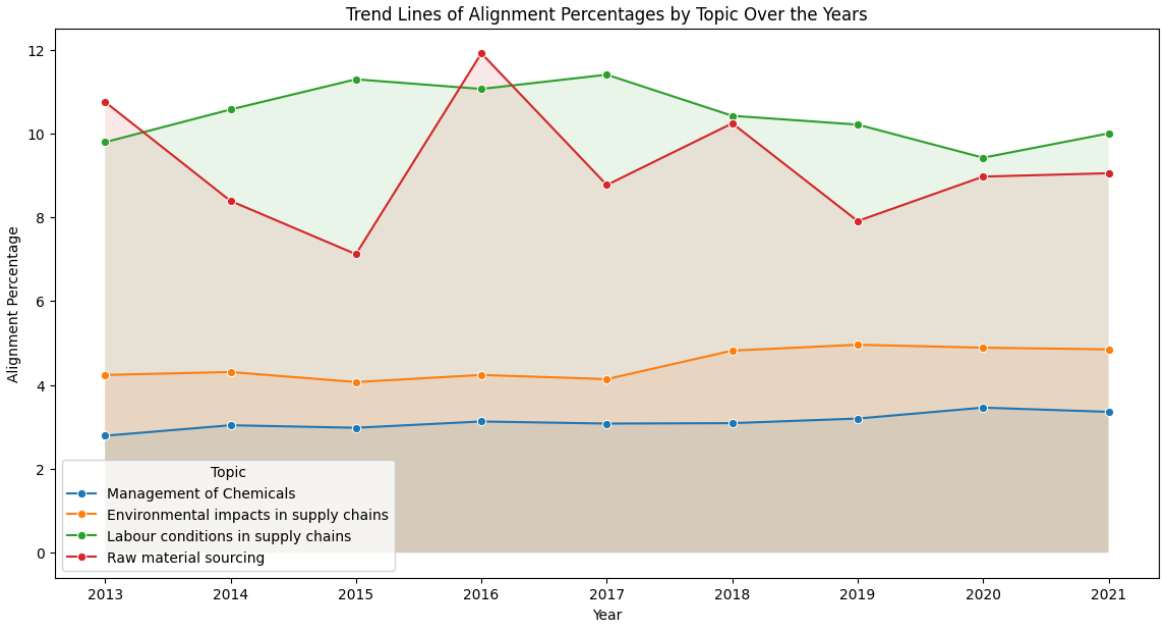
There has been notable variability in attention to raw material sourcing, with scores fluctuating between 7.12% and 11.91%. The variability suggests that the focus on raw material sourcing is influenced by external factors such as regulatory changes or consumer demand for sustainable products. The peaks in certain years could correlate with prominent industry events or shifts in regulatory landscapes that emphasize the importance of sustainable sourcing practices.

Overall, compared to the company-level analysis, the aggregate data highlights a stronger alignment with **raw material sourcing and labour conditions (figure 10)**. This could be attributed to the broader scope of the analysis, which includes a diverse array of companies with varying strategic priorities. Moreover, the relatively lower scores for chemical management both at the company and industry level suggest that this area may not yet be perceived as a critical priority across the broader industry in comparison with other material topics.

Table 7: AP over the years at the industry level

SASB Material Topics / Year (number of SRs analyzed per year)	2013 (10 SRs)	2014 (15 SRs)	2015 (20 SRs)	2016 (25 SRs)	2017 (30 SRS)	2018 (35 SRS)	2019 (40 SRs)	2020 (45 SRs)	2021 (50 SRs)
Management of Chemicals	2.79%	3.04%	2.98%	3.13%	3.08%	3.09%	3.2%	3.46%	3.36%
Environmental impacts on supply chains	4.24%	4.31%	4.07%	4.24%	4.14%	4.82%	4.96%	4.89%	4.85%
Labour conditions in supply chains	9.79%	10.57%	11.29%	11.06%	11.4%	10.42%	10.21%	9.42%	10%
Raw material sourcing	10.75%	8.39%	7.12%	11.91%	8.77%	10.24%	7.91%	8.97%	9.05%

Figure 10: Trend Lines of AP by topic over the years at the aggregate level



5.2 Balance at the industry level

Building upon our comprehensive analytical framework, we extend our purview to incorporate a more expansive range of players within the fashion industry. This expanded analysis scrutinizes SRs across the sector, covering entities of apparel industry counterparts. The objective of this aggregate-level analysis is to penetrate beyond the surface-level narratives of sustainability, offering a rigorous critique of the industry's substantive engagement with sustainability concerns. Our deep dive into the sustainability discourse involves a systematic extraction of the most salient positive and negative

statements from SRs on an annual basis, relevant to each SASB material topic leveraging analytical methods previously applied in the company-specific evaluations. The "top" positive and negative sentences are discerned through a coding process that gauges the intensity of the sentiment they convey. These are not merely sentences with a passing positive or negative tone but rather those that score at the extreme ends of sentiment analysis tools—highlighting the most pronounced emotional expressions that underpin the reports' sustainability narratives. Such a methodical selection enables us to map out and scrutinize the most compelling sustainability claims and acknowledgments of challenges.

This complex approach unveils a more complete comprehensive picture of the industry's transparency and accountability in the post-Rana Plaza era beyond the representatives of the Accord and Alliance. By tracing these polar sentiments over time, we gain valuable insights into the apparel industry's dynamic approach to sustainability.

5.2.1 Management of Chemicals in Products:

The management of chemicals in products within the apparel industry is frequently portrayed through a lens of active engagement and mitigation, even when addressing setbacks. This narrative strategy is likely designed to strengthen credibility and reassure stakeholders following historical critiques or incidents. The language used in these reports suggests an industry effort to project a responsible stance on chemical management.

- **Analysis and Implications**

  The predominance of positive narratives and the constructive framing of negative aspects indicate an industry-wide strategy to project a responsible image in chemical management. This approach could be aimed at enhancing legitimacy and public trust, particularly in response to past criticisms or disasters in the industry.

  Positive narratives focus on proactive measures such as compliance with safety standards, implementation of innovative risk management systems, and the attainment of certifications that demonstrate a commitment to chemical safety and environmental responsibility. These accounts stress the industry’s dedication to maintaining high safety standards, adhering to regulations, and striving for excellence in managing chemical hazards. There is a clear intent to portray the management of chemicals as a controlled, transparent, and continuously improving process, often highlighting certifications, awards, and strategic plans as evidence of these efforts.

Negative disclosures, while acknowledging risks, incidents, or compliance challenges, consistently pivot to emphasize remedial actions, ongoing improvements, or future plans to address these issues. This strategic framing aims to soften the impact of negative information by immediately associating it with positive action or intent. This observation aligns with the concept of "impression management" in corporate reporting, where companies strategically present information to shape stakeholders' perceptions (Merkl-Davies & Brennan, 2008).

While this strategy can reassure stakeholders, it may also lead to skepticism if stakeholders perceive it as deflecting from underlying issues. For example, Laufer (2003) highlights that excessive focus on future plans can be perceived as "greenwashing," where the emphasis on positive intent overshadows the actual severity of current issues.

Negative sentences, even when discussing possible hazards or incidents, tend to emphasize the robustness of the industry's risk management frameworks. These sentences often detail how risks are systematically identified, assessed, and mitigated, reinforcing a narrative of control and responsibility. The emphasis on risk management aligns with findings that companies often highlight their governance and risk management processes to build trust and demonstrate accountability (Beasley et al., 2005).

The industry's reporting on chemical management often frames challenges as opportunities for improvement and learning. This constructive approach in the narrative can reflect a strategic choice to portray the industry as dynamic and responsive to change rather than static or defensive. This aligns with the concept of "transformational leadership" in sustainability reporting, where companies position themselves as proactive and forward-thinking (Eccles & Krzus, 2010).

Table 8: Management of chemical sentences

Examples of negative sentences	Examples of positive sentences

"with respect to the cardinal direction “nature” the focus lies on the reduction of negative impacts: be it the product per se that has a negative impact (e.g., gas burning stoves) or the way our operation works (e.g., lighting in stores; transportation of goods), or the input materials we source (organic or conventional cotton; recycled or virgin polyester)."	"In an effort to manage safety when transporting hazardous substances, Toray concludes security agreements with certain customers, raw material manufacturers, and shipping companies to designate their specific safety responsibilities and roles with regards to safety."
"Processing raw materials such as cotton is a part of the value chain that is often associated with concerns for working conditions and intense water and chemical use".	"Since 2015, all of our denim production is covered by a tool developed by the sustainable denim experts at Jeanologia. This helps us measure, set targets and reduce energy and water use for our wash processes"
"New orders are not placed until the violation has been corrected and the victim has been compensated. "	"- responsible use of innovation and technology promoting responsible and safe use of technology to create new and modified products, packaging and ways of conducting business; considering the impacts of automation."
"it is supported by our group safety and well-being strategic plan and safety management system, which reflects a proactive risk and behaviour-based approach. "	"in an effort to manage safety when transporting hazardous substances, toray concludes security agreements with certain customers, raw material manufacturers, and shipping companies to designate their specific safety responsibilities and roles with regards to safety."
"..will take all necessary steps to prevent accidents and injury arising out of, linked with or occurring in the course of work, or as a result of facility operations. "	"in the “application system development and maintenance safety management regulations” and “system access management regulations” that we formulated for the information safety management system, there are clear requirements for safety and security. "

5.2.2 Environmental impacts in the supply chain

The narrative on environmental impacts within supply chains, as revealed by the analysis of SRs, is painted in a decidedly proactive and positive tone. Positive sentences consistently depict an image of active environmental responsibility and innovation, highlighting efforts in areas like energy consumption, water usage, waste management, and emissions reduction. Companies demonstrated proactive measures to assess and refine these aspects, leading to strategic enhancements.

There was a notable emphasis on stakeholder collabouration, with companies engaging with employees, customers, suppliers, and communities to align on environmental priorities, set objectives, and drive sustainable practices. The proactive stance and stakeholder engagement are reflective of the principles advocated by the GRI and the Carbon Disclosure Project (CDP), which encourage transparent and comprehensive reporting on environmental impacts and sustainability efforts (GRI, 2020; CDP, 2020). However, the overly positive framing of environmental initiatives can sometimes mask the reality of the impacts as Laufer (2003) warns of "greenwashing," where companies may overstate their efforts or underreport negative outcomes to maintain a favorable image.

Conversely, negative statements, while recognizing environmental challenges, are consistently couched within a framework of responsive action and improvement. Disclosures of non-compliance or environmental mishaps are swiftly coupled with corrective measures and strategic initiatives, projecting an image of rectitude and dedication to continuous improvement. For instance, discussions about regulatory non-compliance or penalties were typically followed by assurances of rectification and commitment to high safety standards, underscoring an intention to preserve a good reputation and trust among stakeholders.

This strategic presentation of information seems to serve a dual purpose: it assures stakeholders of the industry's proactive environmental agenda while also shaping a narrative that deflects from the negative impacts. Statements that acknowledge issues such as waste production or inefficiencies in manufacturing are quickly followed by accounts of comprehensive waste management systems or innovative waste reduction measures. The industry's discourse is crafted to underscore a commitment to environmental sustainability, spotlighting the positive while subtly mitigating the negative.

Studies on corporate sustainability reporting frequently identify this dual narrative strategy. For instance, the concept of greenwashing highlights how companies report on negative impacts in a way that emphasizes subsequent improvements, which can lead to an overall positive perception despite ongoing issues (Karnani, 2010). This contradicts research indicating that the environmental footprint of the fashion industry remains considerable, suggesting that despite reported improvements, substantial challenges persist. The Fashion Transparency Index (2021) also highlights a gap between reported initiatives and actual performance, indicating a need for more rigorous accountability measures.

Table below is a tabulated representation of thematic findings, illustrating positive and negative sentences that span from 2015 to 2021, showcasing the industry’s dialogue around environmental impacts:

Table 9: Environmental impacts sentences

Negative sentences	**Positive sentences**
“…Many of our factory partners do not yet have reliable systems to measure and track energy data"	"The reduction was achieved due to productivity improvements combined with steady implementation of

	energy conservation initiatives at plants, as well as efforts by energy efficiency specialists at the company. "
"..risks identified through the sustainability materiality assessment are incorporated into guess inc.'s overall enterprise risk management approach and are subject to internal audit procedures, depending on the level of risk assessed . "	"Led energy savings: Macy's greatest opportunity for energy reduction continues to be our lighting."
"…in addition to this operational waste stream analysis, we are taking action on reducing waste across other aspects of our business. "	" We have a role to play in ensuring we use the best possible mix of energy sources, improve the energy efficiency of our manufacturing and operational processes and reduce the potential climate impact of the products we design and sell. "
"If not properly managed, the resulting pollution can damage habitat and affect the local communities"	"Given that the group's facilities are highly diversified, in order to be able to effectively intervene and reduce energy consumption, we take into consideration the differences of our facilities and we implement special measures and practices for improving their energy efficiency. "
"As these pose risks to our industry and our environment, we believe that addressing them in a meaningful way falls within our responsibility as a business"	"to understand and help reduce our suppliers' environmental footprint, we focus on areas where we can have the largest impact and influence: • responsible sourcing • water stewardship• supplier environmental assessments responsible sourcing we make our products from a variety of natural and man-made materials"

5.2.3 Labour Conditions in supply chains

In the context of labour conditions within supply chains, the narrative crafted in sustainability reports consistently aims to project a conscientious image of corporate responsibility toward workforce welfare. Positive sentences emphasize the commitment to fair labour practices, health and safety, and respect for human rights, suggesting a proactive approach to managing labour conditions and fostering a safe and inclusive work environment.

Critics argue that these narratives often downplay the severity of labour issues, with research indicating that forced labour, exploitation, and unsafe working conditions remain pervasive in global supply chains (Crane et al., 2019). This suggests that while companies report positive actions, the actual conditions may not always reflect these improvements.

Negative statements offer a candid—if somewhat moderated—picture of labour challenges. Although they touch upon serious issues like forced labour, exploitation, and safety hazards, these acknowledgments are frequently tempered by accompanying commitments to improvement. This narrative technique signals an acknowledgment of labour-related problems compared with an emphasis on resolution and reform measures, thereby framing the negatives within a context of constructive action and corporate accountability.

The framing of negative disclosures alongside commitments to improvement is a common strategy in corporate reporting, aimed at maintaining a reputable image while addressing issues (Karnani, 2010).

Independent assessments, however, often reveal gaps between reported commitments and actual practices. Studies have shown that despite reporting improvements, considerable labour violations persist in many supply chains (Locke et al., 2009). This highlights the need for external verification to ensure that reported improvements are genuinely implemented.

For instance, sentences mentioning forced labour or workplace accidents are often paired with descriptions of training programs, policy implementations, or grievance mechanisms designed to prevent such issues in the future. This suggests that while companies are somewhat transparent about the challenges, there is a tendency to frame these negatives in a way that still casts the company in a positive light, possibly to maintain a good reputation.

The table below is a tabulated representation of thematic findings, illustrating positive and negative sentences that span from 2015 to 2021, showcasing the industry's dialogue around labour conditions:

Table 10: labour condition sentences

Negative sentences	Positive sentences
"The program tackles the risks posed by human trafficking, forced labour or debt servitude where coercion, threats or deception can be used to intimidate, penalize or deceive workers, thereby creating situations of involuntary work and exploitation."	"We indirectly create employment for over a million people, not least women, in the countries that manufacture our products"
"We informed the ceo of the supplier that the abuse alleged in the report would not be tolerated and required remediation regarding employment contracts, reporting systems and grievances raised by workers. "	"Going forward, we plan to leverage the SAC facility social labour module (FSLM), which allows facilities to assess the social and labour conditions for its workers, helping ensure they're creating safe and fair working conditions - in an effort to support industry alignment on this important topic."
"we recognise that modern slavery risks exist in the jurisdictions from which our crew are recruited and are particularly conscious of the risks of human trafficking "	"we believe that offering learning and development opportunities will help to ensure our employees feel supported and equipped to carry out their role to the best of their ability. "
"we identify and manage any human rights risks in our operations and supply chain using the following steps: —due diligence. we systematically conduct due diligence to identify, address and report on human rights-related risks or impacts during relevant assessment processes"	" Due to the potential impact of an accident not only within the company but on surrounding communities, the group places the highest priority on accident prevention. "
"possible risks associated with incidents of discrimination or violation of human rights in the workplace might be caused either by the lack of an appropriate management framework by the group's management or its suppliers and partners as well as by the possible inadequate implementation of the appropriate practices by the employees. "	"The creation and safeguarding of employment positions, the occupational health conditions, the meritocracy and personal development, the respect for human rights, as well as the provision of equal opportunities at training, assessment and reward for all, constitute the focus of our group's philosophy and practices. "

5.2.4 Raw Material Sourcing

The extracted sentences for raw material sourcing across the years reveal a strong emphasis on positive statements related to sustainability commitments, ethical standards, and responsible sourcing. Companies shine a spotlight on their robust commitments to environmentally conscious

material choices, such as utilizing certified down and eco-friendly cotton. They frequently cite participation in notable sustainability programs, championing their strides in responsible sourcing and the integration of sustainable materials.

This focus on positive aspects aligns with the increasing demand for transparency and ethical practices in raw material sourcing, as advocated by organizations like the Better Cotton Initiative (BCI) and the Forest Stewardship Council (FSC).

However, the negative sentiments that surface tend to be mild, pointing out hurdles and constraints in the journey towards enhanced sustainability, rather than problems or dire failures in raw material sourcing. Moreover, compared to other topics raw material sourcing seems to have fewer extracted negative sentences. This observation suggests a portrayal of raw material sourcing as a progressive and evolving journey, characterized more by aspirations and ongoing efforts.

The relative scarcity of negative disclosures compared to other topics may imply that the challenges in raw material sourcing are either less pronounced, less openly acknowledged, or simply not as starkly reported. This could suggest a tendency to frame raw material sourcing issues in a more positive light, likely to maintain a good reputation and emphasize a commitment to progress rather than dwelling on problems. This trend is consistent with findings that companies often frame their sustainability efforts positively while downplaying ongoing challenges (Karnani, 2010).

The table below is a tabulated representation of thematic findings, illustrating positive and negative sentences, and showcasing the industry's dialogue around raw material sourcing:

Table 11: Raw Material Sourcing Sentences.

Negative sentences	Positive sentences
"…This process relies on chemicals to prepare the leather, and then those chemicals must be processed using large amounts of water"	"Down beginning with Columbia's fall 2016 global product line, Columbia is using 100% responsible down standard certified natural down based on apparel exchange's responsible down standard. "
" our restricted substances list (rsl) applies to owned-brand clothing and non-clothing apparels. "	"Cotton represents nearly 70% of raw material use, so we have a great need and opportunity to invest in sourcing more sustainable cotton. "
After completing our first round of effluent testing at our key non-denim mills, we plan to extend the program to our denim mills. We also learned that two of our denim producers that supply 14% of denim for Guess US and Guess Canada, or 4% of total production, are already implementing environmental improvements, and is an opportunity area for improvement that we are currently exploring	"apart from using and creating demand for the more sustainable fibre options available on the market, we are part of networks such as mistra future fashion and engaged in development projects that work to find new and resource efficient ways to produce apparel fibres. "
As the availability of non-renewable resources constantly declines, resulting in increasing material costs, we must search for alternative solutions now this means that we need to replace the non-renewable materials we use, like synthetic fibers such as polyester, but also cotton and various blends	we became members of the apparel exchange (te) and we intend to use their responsible wool standard (rws) wool in our products moving forward in order to support our commitment to this goal.

"Less than 1% of the material used to produce clothing is recycled into new clothing, equivalent to a loss of more than $100 billion of materials a year. "	"we have a series of ambitious targets to achieve this aim: • ensure all key materials are 100% traceable by 2025, supported by our use of certified materials where the country of origin is verified and disclosed. "

When scrutinizing the sentences of the reports, we observed that positive sentiments dominate even in negative sentiments that are supposed to highlight the risks and areas for improvement. This positive skew in reporting suggests that even when negative aspects are acknowledged, they are framed within a context of proactive measures and improvements. This pattern raises concerns about the authenticity of the reports and the possibility for greenwashing, where companies may be presenting an overly favorable image to attract and reassure investors.

Considering the literature indicates ongoing challenges in regions like Bangladesh as a result of misconducts of apparel industry contradicts the consistently optimistic portrayal in SRs and highlights critical insights into the communicative strategies employed by these companies. This discrepancy underscores a serious red flag: the positive tone, while expected, maybe more symbolic than substantive, masking the reality of sustainability efforts.

Key Takeaways and Contributions to the Literature:

This robust methodology can greatly assist in analyzing extensive SRs, ensuring that both thematic relevance and sentiment are captured accurately. offering prominent benefits for stakeholders seeking to enhance transparency and accountability in corporate reporting. By leveraging advanced sentiment analysis tools and thematic extraction techniques, this approach sets a new standard for evaluating the comprehensive performance of companies in the realm of sustainability. Future work could involve fine-tuning advanced models with targeted training sets to further improve accuracy and applicability across diverse datasets.

This study's systematic analysis of SRs across the apparel industry reveals several critical insights and contributions. By employing advanced text mining and sentiment analysis, the research identifies a pervasive tendency for companies to focus on positive narratives, often presenting challenges as opportunities. However, this approach lacks specificity in addressing tangible risks or challenges As a result, it becomes harder for readers to understand the real issues companies are facing, potentially obscuring the depth and complexity of these challenges in favour of projecting an image of continuous improvement and responsible practices. The study also finds that companies are increasingly aware of and integrating key materiality topics into their reports, particularly in recent years. However, while these topics are covered, there remains a strong tendency to portray them in an

overly positive light. This discrepancy raises concerns about the likelihood of for greenwashing, especially given that some on-the-ground facts conflict with the positive portrayals in these reports.

This methodological innovation not only highlights the industry's commitment to advancing sustainability practices but also necessitates a critical examination to discern genuine progress from superficial reporting. The study's innovative approach contributes considerably to the literature by providing a robust template for future research in theme extraction and sentiment analysis, moving beyond surface-level analyses to offer deeper insights into the sustainability narratives of companies.

Additionally, the research addresses a critical gap by focusing on both materiality and balanced reporting, offering a nuanced understanding of how sustainability issues are communicated. This comprehensive analysis at the industry level extends the purview beyond individual company narratives, providing valuable insights into the apparel industry transparency and accountability post-Rana Plaza. The findings serve as a critical resource for stakeholders, enabling a more informed assessment of the credibility of sustainability claims and the true extent of corporate commitment to sustainability. By systematically dissecting the narratives within SRs, this study enhances the understanding of sector-wide sustainability practices, contributing to the discourse on credible and truthful reporting, and setting a precedent for future research to scrutinize the gap between disclosures and actual performance.

6 Discussion

In this section, we provide a high-level summary and synthesis of our research objectives and findings, highlighting key insights from the analysis in the previous section. This is followed by a discussion of the contributions our study makes to both research and practice, highlighting how this work advances the academic understanding of SR, from both methodological and analytical perspectives, while also offering practical implications for companies and policymakers.

6.1 High-level summary of findings:

Our study tackles several important but underexplored areas in current SR research. First, existing studies often focus exclusively on either financial materiality—how sustainability issues impact a company’s financial performance—or sustainability materiality, which considers the broader societal and environmental impacts. Yet both practice and governance are increasingly moving towards integrated reporting (often referred to as double materiality) that identifies a core set of themes of relevance to multiple stakeholder groups including the finance, civil society, and regulators. Second, there is a lack of research that examines both whether the information contained in SRs is material

(the “what”) and balanced (the “how”), meaning whether companies are presenting a fair and accurate portrayal of both positive and negative outcomes on topics and issues of relevance to the audience that is using these reports. Finally, much of the existing research tends to focus on a limited number of SR reports that cover a small number of companies and short timeframes. This restricts the validity and generalizability of findings and the insights that can be garnered from such analysis.

We address these issues in several ways. First, we conduct a materiality analysis that integrates key elements from both the GRI and SASB frameworks. This integration allows us to capture the core aspects of materiality (both financial and sustainability) that are important to multiple stakeholder communities and unarguably essential to understanding what material to the apparel industry is. This provides a holistic and immutable view of what is truly significant for both companies and society, allowing us to apply this framework retrospectively to SR data from the past.

Second, we complement this integrated materiality analysis with a detailed examination of balance in SR. This involves assessing whether companies are offering a realistic narrative that acknowledges both positive (progress, issue resolution, successes) and negative (challenges, delays, failures, etc.), accurately reflecting their sustainability efforts and shortcomings. Doing so is crucial because it ensures that the sustainability reports provide a more comprehensive and truthful depiction of a company's performance, which is essential for building stakeholder trust and enabling informed decision-making by investors, regulators, and other interested parties.

Third, we significantly expand the scope of analysis to include hundreds of SRs spanning multiple years, at both company-specific and industry-wide levels. This significantly enhances the robustness and generalizability of our findings and allows us to identify broader trends, patterns, and potential anomalies in sustainability reporting practices. By doing so, we provide deeper insights into how sustainability narratives evolve over time and across different contexts, thereby contributing to a more comprehensive understanding of the SR reporting.

These three objectives, composing a multi-faceted analysis of SR, are supported by the application of advanced analytical tools and techniques. Some, of these are recently available and therefore have seen limited adoption, particularly within the field of SR analysis. These tools enable us to deliver a more comprehensive evaluation of SR – analysis of large datasets that simultaneously increases both the depth and breadth of analysis (evaluating both materiality and balance across numerous companies and at the industry level, across nearly a decade of reporting).

Our detailed company-level analysis reveals several key insights. First, our findings show that companies' SR is consistently focused on what are considered to be the main material topics during the time period that was analyzed – these material topics were identified via our heatmap analysis of GRI and SASB. This finding is of interest because holds true during the entire time period. When understood in terms of the early years in the period, this suggests that there may be an immutable core of material topics that were readily known and identifiable by companies in the apparel industry, even prior to key standards such as GRI being widely adopted. When understood in terms of the latter years, it in turn suggests that frameworks like GRI (the most widely adopted in the latter years of the sample, before developments of SASB), in codifying the material topics.

Second, our thematic alignment analysis did however uncover an unexpected finding in terms of the theme that is consistently prioritized at the company level. Many companies, especially those within the Accord and Alliance groups, show a stronger alignment with environmental than labour issues. This occurs in most of the years after the Rana Plaza disaster, expect for one – 2014, the year immediately following the disaster itself. This is surprising because despite significant pressures to increase both the quality and volume of reporting on labour issues and the stated commitments of the companies themselves to do so, the companies that were most affected by Rana Plaza and signed up to the Accord and Alliance, consistently prioritized environmental topics in their SRs. Even more surprising is that the analysis at the industry level (which we summarize below) does not have environmental theme coming out on top as a leading priority.

A third notable finding is that the materiality analysis which focused on alignment percentage (AP) trends reveals considerable variability across different companies and over time, with a general upward trend in AP across the years. Over time, different sustainability themes—such as environmental impacts, labor conditions, and raw material sourcing—have shifted in priority, with some themes receiving more focus in certain years than others. For example, in 2016, H&M placed a higher emphasis on labor conditions, while environmental impacts dominated their reporting in other years. Similarly, Next demonstrated fluctuations in their focus, with labor conditions being a priority in the early years, followed by a shift towards environmental concerns in subsequent years.

At the industry level, this variation in thematic focus becomes even more pronounced and surprising, particularly when compared to the company-level focus. Labour conditions and raw material sourcing often took precedence, with raw material sourcing sometimes emerging as the top priority, even surpassing labor and environmental issues. Environmental impacts, while still significant, generally ranked third in terms of priority. Management of chemical topics, however, consistently

exhibits lower alignment percentages compared to environmental impacts, labor conditions, and raw material sourcing. The relatively lower scores for chemical management at both the company and industry levels suggest that this area may not yet be perceived as a critical priority across the broader industry, especially when compared to other material topics.

A noticeable trend is also reflected in the balance scores. At the company level, the rise in balance scores indicates a growing proportion of positive sentences compared to negative ones, suggesting that companies are increasingly emphasizing their achievements in sustainability reporting. At the industry level, even when negative sentences are present, they often downplay risks and challenges, leaning towards a more positive framing. This indicates that, as AP increases over time, the content also becomes more positively skewed across these themes, emphasizing favourable outcomes while underreporting challenges. This shift towards more positive reporting is evident both in the frequency of alignment with sustainability themes and in the overall sentiment of the reports.

Fourth, we observe consistently high levels of positivity in these reports (both at the company and industry level), which is overshadowed by a large margin discussion of challenges and areas requiring improvement. While it is not unexpected that sustainability reports would highlight positive achievements, the extent to which this positivity dominates the narratives is surprising. This overwhelming focus on positive aspects, with little mention of challenges or areas needing improvement, or framing of issues in positive terms, is concerning as it signals the likelihood of greenwashing communication strategies, where companies exaggerate their environmental or social achievements to create a misleadingly favourable image. Such a pattern suggests that these reports are likely not providing transparent or accurate accounts of a company's true sustainability performance, thus potentially masking and putting an overly positive spin on significant issues that need to be addressed. This undermines the value of the information provided in SRs and may lead to the loss of stakeholder trust – investors, civil society, and regulators might struggle to discern actual progress and identify areas needing improvement, potentially leading to skepticism about the reported sustainability efforts – an issue whose salience has grown significantly in recent years. This third main finding is particularly surprising and concerning given the research context itself - an industry under significant pressure to be more transparent and one that seemingly has a window of opportunity in the years following Rana Plaza to acknowledge the challenges they are facing, particularly when it comes to labour issues, in a more balanced and open manner.

6.2 Contributions

6.2.1 Contributions to the literature on SR reporting

Our methodology represents a considerable advancement in SR analysis. Until now, most research has relied on either manual approaches or text mining techniques that were limited in various ways—either by the tools themselves, which often lacked sufficient depth or breadth, or by the scope of the analysis. Manual content analysis, while often detailed, was constrained by the sheer volume of data, making it challenging to scale or apply consistently across large datasets. Researchers often struggled to maintain objectivity and consistency when manually coding and interpreting large amounts of text, which inevitably led to limitations in the depth of the insights gained. Due to the laborious coding process involved, they also struggled to deal with the expansive corpus of SR texts and information, unable to take full advantage of the growing amount of data available for analysis.

On the other hand, automated text mining techniques, though an improvement over manual methods, have to date been applied in a restricted manner that also constrained their potential insights (see paper 1 of this thesis). For example, tools such as word counts and basic frequency analysis have provided some interesting insights, such as identifying common themes or key terms. However, these approaches are limited in their ability to capture the nuanced meanings and contextual subtleties within the texts. Additionally, text mining techniques have often been applied to relatively small datasets, composed of a few companies and limited timeframes, which restricted the breadth of analysis and the generalizability of the findings.

In contrast, our approach in this study employs advanced text mining and natural language processing (NLP) techniques to provide much deeper and broader insights into the phenomenon of SR. We have developed and implemented a method that enables large-scale, automated, accurate, and reliable analysis of sustainability reports. This method systematically extracts key themes, categorizes sentences into those highlighting opportunities and strengths versus risks and areas for improvement, and aligns these findings with material topics identified by the SASB and GRI frameworks. By utilizing the BERT model for encoding and classifying sentences and VADER's sentiment analysis tool, we offer a detailed assessment of balance within SRs.

This approach not only provides a robust framework that can be applied at both the company and industry levels but is also applicable over extended periods. Our goal is to set a new benchmark in SR analysis, one that is replicable by practitioners and analysts alike, enabling them to conduct

comprehensive, reliable, and insightful evaluations of sustainability communications across various contexts.

Analytical contributions:

This study also makes several key contributions to the field of SR by addressing gaps in existing research and advancing our understanding of how companies in the apparel industry prioritize and report on key sustainability issues.

First, our analysis challenges the prevailing assumption that labor issues would dominate sustainability reports (SRs) in the years following the Rana Plaza disaster, especially within companies that are part of the Accord and Alliance groups. Contrary to expectations, we found that these companies exhibited a stronger alignment with environmental issues than with labor concerns. This skewed thematic alignment is particularly surprising given the intense social and regulatory pressures to improve labour conditions in the apparel industry during this period. Perhaps most surprising of all is the fact that the broader industry did not exhibit this same pattern – rather, the environmental theme was often second or third priority. This insight adds nuance to our understanding of corporate behaviour in the face of regulatory and social scrutiny, raises questions about the impact of multi-stakeholder initiatives like the Accord and Alliance, and raises questions about the overall sincerity and transparency of SR in the industry.

We acknowledge that this surprising finding may however be subject to multiple, differing interpretations. On the one hand, this is unexpected given the intense social and regulatory pressures regarding labour issues that the companies that were most affected were under in the years following the Rana Plaza disaster - one would expect the highest priority theme to be related to labour issues, at least in the years following Rana Plaza with some tapering off as the issue salience subsided. Moreover, we would also expect the companies involved in the Accord and Alliance to prioritize labour issues over companies that are not members of the same initiatives. On the other hand, this may be interpreted as a signal that there may have been an understanding and awareness on the part of companies that labour issues in a place like Bangladesh were difficult to resolve and that there was a need to “turn the page” and/ or emphasize other areas where more positive developments may have been taking place. Unfortunately, the literature analyzing post-Rana Plaza sustainability performance has little insight to offer on this point as it has focused almost exclusively on labour issues.

Second, our research opens up a new perspective on the notion of the temporal dynamics of SR. We uncover significant variability in alignment percentages (AP) across companies and over time. While

we observed a general upward trend in AP, indicating a growing alignment with key sustainability topics, the often considerable fluctuations at both the company and industry levels reveal how companies shift their focus between different sustainability themes over time, in some cases from year to year which is a fairly short time period. For instance, our findings show that companies like H&M and Next have varied their emphasis on labour conditions and environmental impacts across different years. It is not clear what is driving such variability - changes in internal priorities, external pressures, a combination of these, or something else. Additionally, the consistently lower alignment percentages for chemical management topics over the entire length of our sample suggest that certain issues may not be receiving the attention they warrant within the broader industry. As such, by highlighting the temporal dimension, our study contributes to a more dynamic understanding of how companies navigate and prioritize competing sustainability concerns over time, an important aspect that had been missing in the SR research to date.

Third, this study contributes to the literature on greenwashing. Our findings reveal a concerning trend toward consistently positive reporting, where the emphasis on achievements and successes often significantly overshadows discussion of challenges and areas requiring improvement. At both the company and industry levels, the significant rise in balance scores, which indicate a higher proportion of positive statements, suggests that companies are crafting narratives that downplay risks and difficulties to a worrisome degree. This trend toward much more positive framing points to the possibility of greenwashing, where companies may be purposefully exaggerating their sustainability achievements to create a more favourable public image. This finding is particularly critical as it underscores the lack of transparency and balance in SRs, which jeopardizes stakeholder trust but also fails to ensure that investors, regulators, and civil society organizations have access to accurate and comprehensive information on corporate sustainability performance, a point we turn to below.

Overall, this study contributes to the ongoing debate on the credibility and reliability of SR by providing empirical evidence of the limitations of current reporting practices. By developing a new approach to examining the issue – involving thematic alignment, prioritization, and balance in SRs across a broad sample of companies and years - we offer a more comprehensive and nuanced understanding of how sustainability narratives are communicated within the apparel industry and of the extent of their limitations. The dominance of positive reporting, especially in an industry under significant pressure to improve transparency post-Rana Plaza, highlights the need for more rigorous standards and oversight to prevent the misuse of SR as a tool for public relations rather than a

genuine reflection of corporate sustainability efforts and performance, which brings us to our next contribution.

6.2.2 Contributions to SR practice and governance

These insights are valuable not only for academic researchers but also for practitioners, analysts, policymakers, and regulators who are working to improve the standards and practices of sustainability reporting worldwide. Our research provides a robust analytical framework that reveals important issues, and as such may be used to both detect problematic SR practices as well as improve the accuracy and transparency of corporate disclosures.

The findings from our materiality analysis and balance assessment are relevant for company managers responsible for sustainability reporting. The type of analysis performed here can be adopted and applied by the creators of the reports – the internal analysts, communications strategists, and other personnel crafting the reports – to better manage content disclosure and mitigate the growing risks associated with the practice of greenwashing, whether done in an intentional manner or not. In an era where companies face increasing scrutiny from investors, regulators, and civil society, maintaining credibility through transparent reporting is essential. Our analysis, which integrates both the GRI and SASB frameworks to assess materiality and assess whether disclosures are balanced, shows that companies must focus on genuinely significant sustainability issues that matter not only to their financial performance but also to broader societal and environmental impacts. The advanced methodological approaches we employed can identify potential greenwashing practices, offering a critical tool for companies to ensure the integrity of their sustainability reports. Some consulting firms, such as KPMG, are already experimenting with NLP tools to assess balanced reporting, and our study contributes to these pioneering efforts by providing a comprehensive methodological and technological framework that can be adopted and improved. Overall, this approach may help companies avoid the pitfalls of greenwashing and safeguard their reputation, by ensuring that their reports are meaningful and transparent.

Furthermore, we believe our research serves as a valuable model for governments, civil society organizations, and regulatory bodies in scrutinizing corporate disclosures in SRs. The detailed, integrated examination of both materiality and balance in our analysis - where we assess whether companies present a realistic narrative that reports on the right things and acknowledges both achievements and shortcomings - provides governments and NGOs with a tool that can assist them in holding companies accountable. Governments, in particular, are increasingly interested in ensuring

that SRs are both material and balanced because these reports inform policy and regulatory decisions, guide investment strategies, and shape public perceptions of corporate responsibility. By ensuring that SRs accurately reflect a company's sustainability efforts, including both successes and challenges, regulators can foster greater corporate accountability and drive more meaningful progress toward environmental and social goals. Moreover, material and balanced SRs are crucial for avoiding misleading representations that could result in poorly informed policy decisions or investments. By adopting the advanced analytical techniques used in our study, such as BERT and VADER for sentiment and thematic analysis, governments and NGOs, or the service providers they contract to do so, can better monitor and enforce compliance with sustainability standards, ensuring that companies do not merely pay lip service to ESG commitments but genuinely work toward achieving measurable and impactful outcomes.

Turning to the specific context of the apparel industry in post-Rana Plaza Bangladesh, our findings reveal a significant gap between sustainability rhetoric and actual performance. Despite the industry's increased focus on sustainability reporting transparency and significant pressure to acknowledge and address the barriers to improving labour conditions in particular, our analysis uncovers consistently high levels of positivity in these reports, often overshadowing discussions of challenges and areas requiring improvement. This pattern, which suggests a potential greenwashing strategy, is particularly surprising and concerning given the research context—an industry under significant pressure to be more transparent and one that had a clear window of opportunity, especially in the years following Rana Plaza, to acknowledge the substantial challenges it faces, particularly in labor practices, in a more balanced and open manner. This research underscores the need for continuous improvement in both corporate practices and the regulatory frameworks that govern them, advocating for stronger oversight and more rigorous standards to ensure that sustainability efforts are both genuine and impactful.

6.3 Limitations and Future Research

Despite the notable contributions of this study, several limitations should be acknowledged, which also point to avenues for future research.

First, while our analysis covers a broad sample of SRs across multiple companies and years, it remains focused on the apparel industry. This industry-specific focus, while providing deep insights into a particularly important and high-impact sector, may limit the generalizability of our findings to

other industries. Different sectors may prioritize and report on sustainability issues in ways that differ from the apparel industry. Future research could expand the scope of analysis to include a comparative study across multiple industries, which would provide a more comprehensive understanding of how thematic alignment, balance, and reporting practices vary across sectors.

Second, our study primarily relies on advanced text mining and natural language processing (NLP) techniques to assess the materiality and balance in SRs. While these methods offer significant advantages in terms of scalability and objectivity, they also have limitations. For instance, the particular sentiment analysis tools we used may not fully capture the nuanced ways in which companies frame their sustainability narratives, particularly when dealing with very complex or ambiguous issues that involve political . Future research could explore the integration of more sophisticated NLP techniques that have recently become available such as ChatGPT, LLaMA, and others, for sentiment analysis and text categorization. Such tools were not yet available and robust enough to utilize when this study was being planned and when BERT represented the cutting edge. Additionally, the accuracy of the thematic alignment analysis depends on the comprehensiveness of the frameworks (e.g., GRI, SASB) we employed. Going forward, future studies may consider the inclusion of alternative frameworks that might capture different/ additional dimensions of sustainability reporting.

Third, our analysis is retrospective, examining SRs over a specific period in the past. While this approach allows us to identify trends and shifts in reporting practices over time, it does not account for the evolving nature of sustainability issues or reporting standards that may have emerged more recently. As sustainability reporting continues to develop, particularly with the increasing adoption of integrated and double materiality frameworks, future research should focus on real-time or forward-looking analyses to capture how companies are adapting their reporting practices to meet new standards and expectations. Longitudinal studies that track changes over an extended period, including the latest reports, would also be valuable in understanding the long-term impacts of regulatory changes, societal pressures, and market dynamics on sustainability reporting.

Finally, our study highlights a concerning trend toward a very high positive skewness in SRs, which raises questions about the potential for greenwashing. However, we do not claim that have proof of greenwashing. Rather we are careful to point out that our findings suggest this may be the case or that it may be interpreted as being so. Nor does our analysis directly investigate the motivations behind the trends we have identified or its impact on stakeholder perceptions and decision-making. Future research could delve into the drivers of this significantly positive skewness, exploring factors

such as corporate governance structures, external pressures, industry associations, and market incentives that may encourage companies to present their sustainability performance in an overly favorable light. Additionally, experimental or survey-based research could examine how different stakeholder groups perceive and respond to these very positively skewed reports, providing insights into the effectiveness of reporting strategies and credibility of current sustainability reporting practices.

In conclusion, while this study provides a robust and comprehensive analysis of sustainability reporting practices in the apparel industry, it also opens up several important avenues for future research. Addressing the limitations identified here will be crucial for advancing our understanding of sustainability reporting and ensuring that SRs serve as reliable tools for transparency, accountability, and informed decision-making across all sectors.

7 Conclusion

This study provides a comprehensive examination of sustainability reporting (SR) practices within the apparel industry, particularly focusing on companies implicated in the Rana Plaza disaster. By applying advanced text mining and natural language processing (NLP) techniques, the research uncovers prominent aspects of how these companies report on material topics and maintain balance within their disclosures. Our findings indicate a pervasive tendency towards emphasizing positive aspects of sustainability efforts while underreporting or positively reframing risks and areas for improvement, raising concerns about the credibility of these reports and the possibility of greenwashing.

The study advances the methodology for SR analysis by demonstrating the effectiveness of automated text analysis tools in providing a more nuanced and scalable assessment of corporate disclosures. This approach allows for a more detailed evaluation of thematic alignment and balance, which are critical for ensuring that sustainability reports provide a transparent and accurate reflection of a company's sustainability performance.

Practically, the insights gained from this research can guide companies in enhancing the transparency and balance of their sustainability reporting. By adopting the methods outlined in this study, companies can benchmark their SR practices against industry standards, identify areas for improvement, and ultimately build greater trust with stakeholders.

Overall, this study contributes to the ongoing discourse on sustainability reporting by highlighting the need for more rigorous and transparent reporting practices in the apparel industry. It underscores

the importance of balanced reporting as a cornerstone of credible sustainability communication and provides a robust framework for future research and practical applications in the field.

References

Adam, C. A. (2022). EU v IFRS: Fundamentally different approaches to sustainability reporting. https://drcaroladams.net/eu-v-ifrs-fundamentally-different-approaches-to-sustainability-reporting/

Aggarwal, C. C., & Zhai, C. (2012). A Survey of Text Classification Algorithms. In C. C. Aggarwal & C. Zhai (Eds.), Mining Text Data (pp. 163–222). Springer US. https://doi.org/10.1007/978-1-4614-3223-4_6

Agrawal, R., & Srikant, R. (1994). Fast algorithms for mining association rules. Proc. 20th Int. Conf. Very Large Data Bases, VLDB, 1215, 487–499. https://www.it.uu.se/edu/course/homepage/infoutv/ht08/vldb94_rj.pdf

Aguado-Correa, F., de la Vega-Jiménez, J. J., López-Jiménez, J. M., Padilla-Garrido, N., & Rabadán-Martín, I. (2023). Evaluation of non-financial information and its contribution to advancing the sustainable development goals within the Spanish banking sector. European Research on Management and Business Economics, 29(1), 100211. https://doi.org/10.1016/J.IEDEEN.2022.100211

Albitar, K., Abdoush, T., & Hussainey, K. (2022). Do corporate governance mechanisms and ESG disclosure drive CSR narrative tones? International Journal of Finance & Economics. https://doi.org/10.1002/IJFE.2625

Alliance for Corporate Transparency. (2020). Alliance for Corporate Transparency. https://www.allianceforcorporatetransparency.org/

Al-Shaer, H., & Hussainey, K. (2022). Sustainability reporting beyond the business case and its impact on sustainability performance: UK evidence. Journal of Environmental Management, 311, 114883. https://doi.org/10.1016/J.JENVMAN.2022.114883

Angin, M., Taşdemir, B., Yılmaz, C. A., Demiralp, G., Atay, M., Angin, P., & Dikmener, G. (2022). A RoBERTa Approach for Automated Processing of Sustainability Reports. Sustainability, 14(23), 16139.

Antons, D., Grünwald, E., & Cichy, P. (2020). The application of text mining methods in innovation research: Current state, evolution patterns, and development priorities—Antons—2020—R&D Management—Wiley Online Library. R&D Management, 50(3), 329–351.

Aras, G., & Crowther, D. (2008). Corporate Sustainability Reporting: A Study in Disingenuity? Journal of Business Ethics, 87(1), 279–288. https://doi.org/10.1007/S10551-008-9806-0

Auke, E., & Simaens, A. (2019). Corporate responsibility in the fast fashion industry: How media pressure affected corporate disclosure following the collapse of Rana Plaza. International Journal of Entrepreneurship and Innovation Management, 23(4), 356–382. https://doi.org/10.1504/IJEIM.2019.100800

Aureli, S. (2017). A comparison of content analysis usage and text mining in CSR corporate disclosure. International Journal of Digital Accounting Research, 17, 1–32. https://doi.org/10.4192/1577-8517-V17_1

Aureli, S., Medei, R., Supino, E., & Travaglini, C. (2016). Sustainability disclosure after a crisis: A text mining approach. International Journal of Social Ecology and Sustainable Development, 7(1), 35–49. https://doi.org/10.4018/IJSESD.2016010102

Aureli, S., Medei, R., Supino, E., & Travaglini, C. (2017). Sustainability disclosure and a legitimacy crisis: Insights from two major cruise companies. European Journal of Tourism Research, 17, 149–163.

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

Barbier, L. M., Green, J. L., & Draper, D. S. (2022). The need for open access and natural language processing. Proceedings of the National Academy of Sciences, 119(15), e2200752119. https://doi.org/10.1073/pnas.2200752119

Barcellos de Paula, L., & Gil-Lafuente, A. M. (2021). Algorithm applied in corporate sustainability. Kybernetes, 50(7), 1982–2000.

Bartholomew, D. J., Knott, M., & Moustaki, I. (2011). Latent variable models and factor analysis: A unified approach. John Wiley & Sons.

Baumüller, J., & Sopp, K. (2021). Double materiality and the shift from non-financial to European sustainability reporting: Review, outlook and implications. Journal of Applied Accounting Research, 23(1), 8–28.

Beasley, M. S., Clune, R., & Hermanson, D. R. (2005). Enterprise risk management: An empirical analysis of factors associated with the extent of implementation. Journal of Accounting and Public Policy, 24(6), 521–531. https://doi.org/10.1016/j.jaccpubpol.2005.10.001

Benites-Lazaro, L. L., Giatti, L., & Giarolla, A. (2018). Sustainability and governance of sugarcane ethanol companies in Brazil: Topic modeling analysis of CSR reporting. Journal of Cleaner Production, 197, 583–591. https://doi.org/10.1016/J.JCLEPRO.2018.06.212

Betti, G., Consolandi, C., & Eccles, R. G. (2018). The relationship between investor materiality and the sustainable development goals: A methodological framework. Sustainability, 10(7), 2248.

Billio, M., Costola, M., Hristova, I., Latino, C., & Pelizzon, L. (2021). Inside the ESG ratings: (Dis)agreement and performance. Corporate Social Responsibility and Environmental Management, 28(5), 1426–1445. https://doi.org/10.1002/CSR.2177

Blasco, J. L. (2017). The KPMG Survey of Corporate Responsibility Reporting 2017.

Blazkova, T., Pedersen, E. R. G., Andersen, K. R., & Rosati, F. (2023). Greenwashing debates on Twitter: Stakeholders and critical topics. Journal of Cleaner Production, 427, 139260. https://doi.org/10.1016/j.jclepro.2023.139260

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003a). Latent Dirichlet Allocation. Journal of Machine Learning Research.

Bodendorf, F., Dimitrov, G., & Franke, J. (2022). Analyzing and evaluating supplier carbon footprints in supply networks. Journal of Cleaner Production, 372, 133601. https://doi.org/10.1016/J.JCLEPRO.2022.133601

Boiral, O., & Heras-Saizarbitoria, I. (2020). Sustainability reporting assurance: Creating stakeholder accountability through hyperreality? Journal of Cleaner Production, 243, 118596. https://doi.org/10.1016/J.JCLEPRO.2019.118596

Brammer, S., & Pavelin, S. (2004). Building a good reputation. European Management Journal, 22(6), 704–713.

Breijer, R., & Orij, R. P. (2022). The Comparability of Non-Financial Information: An Exploration of the Impact of the Non-Financial Reporting Directive (NFRD, 2014/95/EU). Accounting in Europe, 19(2), 332–361. https://doi.org/10.1080/17449480.2022.2065645

Busco, C., Consolandi, C., Eccles, R. G., & Sofra, E. (2020). A Preliminary Analysis of SASB Reporting: Disclosure Topics, Financial Relevance, and the Financial Intensity of ESG Materiality. Journal of Applied Corporate Finance, 32(2), 117–125. https://doi.org/10.1111/jacf.12411

Bushee, B. J., Gow, I. D., & Taylor, D. J. (2018). Linguistic Complexity in Firm Disclosures: Obfuscation or Information? Journal of Accounting Research, 56(1), 85–121. https://doi.org/10.1111/1475-679X.12179

Calabres, A., Costa, R., Ghiron, N., & Menichini, T. (2019). Materiality analysis in sustainability reporting: A tool for directing corporate sustainability towards emerging economic, environmental and social opportunities. Technological and Economic Development of Economy, 25, 1–23. https://doi.org/10.3846/tede.2019.10550

Calabrese, A., Costa, R., Ghiron, N. L., & Menichini, T. (2017). Materiality analysis in sustainability reporting: A method for making it work in practice. European Journal of Sustainable Development, 6(3), 439–439.

Calabrese, A., Costa, R., Levialdi, N., & Menichini, T. (2016). A fuzzy analytic hierarchy process method to support materiality assessment in sustainability reporting. Journal of Cleaner Production, 121, 248–264.

Calabrese, A., Costa, R., & Rosati, F. (2015). A feedback-based model for CSR assessment and materiality analysis. Accounting Forum, 39(4), 312–327.

Calabrese, A., Costa, R., Tiburzi, L., & Brem, A. (2023). Merging two revolutions: A human-artificial intelligence method to study how sustainability and Industry 4.0 are intertwined. Technological Forecasting and Social Change, 188, 122265. https://doi.org/10.1016/J.TECHFORE.2022.122265

Caliskan, A., Ozturkoglu, O., & Ozturkoglu, Y. (2022). Ranking of Responsible Automotive Manufacturers According to Sustainability Reports Using PROMETHEE and VIKOR Methods. Advanced Sustainable Systems, 6(6), 2100301. https://doi.org/10.1002/ADSU.202100301

Cerbone, D., & Maroun, W. (2020). Materiality in an integrated reporting setting: Insights using an institutional logics framework. The British Accounting Review, 52(3), 100876. https://doi.org/10.1016/J.BAR.2019.100876

Chang, D.-S., & Cheng, Y.-W. (2013). Explore the effects of industrial context and leaders' viewpoints on corporate sustainability in Taiwan by text mining. Service Systems and Service Management (ICSSSM), 670–673. https://doi.org/10.1109/ICSSSM.2013.6602594

Chen, S., & Bouvain, P. (2009). Is corporate responsibility converging? A comparison of corporate responsibility reporting in the USA, UK, Australia, and Germany. Journal of Business Ethics, 87(SUPPL. 1), 299–317. https://doi.org/10.1007/S10551-008-9794-0/METRICS

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016: 785-794. ACM, New York, NY.

Chiarini Tremblay William, M., Parra, C. M., & Castellanos William, A. (2015). Corporate social responsibility reports: Understanding topics via text mining.

Cho, C. H. (2009). Legitimation Strategies Used in Response to Environmental Disaster: A French Case Study of Total SA's Erika and AZF Incidents. European Accounting Review, 18(1), 33–62. https://doi.org/10.1080/09638180802579616

Cho, C. H., Roberts, R. W., & Patten, D. M. (2010). The language of US corporate environmental disclosure. Accounting, Organizations and Society, 35(4), 431–443. https://doi.org/10.1016/J.AOS.2009.10.002

Clark, K. P., Plc, A., Kotsantonis, S., Rehnberg, C., Ward, B., Advisors, K., Serafeim, G., Indahl, R., Gunvor Jacobsen, H., Equity, S., Davis, T., Lescott, B., Business School, H., Lins, K. V., Bizoumi, T., Lazaridis, S., Stamou, N., Ramani, V., Pinney, C., … Capital Management, R. (2019). Four Things No One Will Tell You About ESG Data. Journal of Applied Corporate Finance, 31(2), 50–58. https://doi.org/10.1111/JACF.12346

Clarkson, P., Li, Y., & Richardson, G. (2008). Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. Accounting, Organizations and Society, 33(4–5), 303–327.

Clarkson, P. M., Ponn, J., Richardson, G. D., Rudzicz, F., Tsang, A., & Wang, J. (2020). A Textual Analysis of US Corporate Social Responsibility Reports. Abacus, 56(1), 1009–1051. https://doi.org/10.1111/ABAC.12182

Collins-Thompson, K. (2014). Computational assessment of text readability: A survey of current and future research. ITL - International Journal of Applied Linguistics, 165(2), 97–135. https://doi.org/10.1075/itl.165.2.01col

Corazza, L., Truant, E., Scagnelli, S. D., & Mio, C. (2020). Sustainability reporting after the Costa Concordia disaster: A multi-theory study on legitimacy, impression management and image restoration. Accounting, Auditing and Accountability Journal, 33(8), 1909–1941. https://doi.org/10.1108/AAAJ-05-2018-3488

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21–27.

Crane, A., LeBaron, G., Allain, J., & Behbahani, L. (2019). Governance gaps in eradicating forced labor: From global to domestic supply chains. Regulation & Governance, 13(1), 86–106. https://doi.org/10.1111/rego.12162

Cui, J., Wang, Z., Ho, S.-B., & Cambria, E. (2023). Survey on sentiment analysis: Evolution of research methods and topics. Artificial Intelligence Review, 56(8), 8469–8510. https://doi.org/10.1007/s10462-022-10386-z

D'Amato, V., D'Ecclesia, R., & Levantesi, S. (2021). Fundamental ratios as predictors of ESG scores: A machine learning approach. Decisions in Economics and Finance, 44(2), 1087–1110. https://doi.org/10.1007/S10203-021-00364-5

Delmas, M. A., & Burbano, V. C. (2011). The Drivers of Greenwashing. California Management Review, 54(1), 64–87. https://doi.org/10.1525/cmr.2011.54.1.64

Deloitte. (2023). A practical approach to assess financial materiality. https://www2.deloitte.com/nl/nl/pages/sustainability/articles/a-practical-approach-to-assess-financial-materiality.html

Deng, Q., Hine, M., Ji, S., & Sur, S. (2017). Building an Environmental Sustainability Dictionary for the IT Industry. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2017.112

Deng, Q., Hine, M., Ji, S., & Sur, S. (2017). Building an Environmental Sustainability Dictionary for the IT Industry. Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2017.112

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805, 1, 4171–4186. https://doi.org/10.48550/arXiv.1810.04805

Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. The Accounting Review, 86(1), 59–100.

DiMaggio, P. (2015). Adapting computational text analysis to social science (and vice versa). Big Data & Society, 2(2), 205395171560290-. https://doi.org/10.1177/2053951715602908

Ding, D., Liu, B., & Chang, M. (2023). Carbon Emissions and TCFD Aligned Climate-Related Information Disclosures. Journal of Business Ethics, 182(4), 967–1001. https://doi.org/10.1007/S10551-022-05292-X/TABLES/16

Du, S., & Yu, K. (2021). Do Corporate Social Responsibility Reports Convey Value Relevant Information? Evidence from Report Readability and Tone. Journal of Business Ethics, 172(2), 253–274. https://doi.org/10.1007/S10551-020-04496-3/TABLES/9

Duan, G., Hu, W., Xue, H., & Lu, X. (2018). Financial risk assessment model of block chain based on soft computing and neural network. IPPTA: Quarterly Journal of Indian Pulp and Paper Technical Association, 30(4), 108–118.

Eccles, R. G., & Krzus, M. P. (2010). One report: Integrated reporting for a sustainable strategy. John Wiley & Sons. https://books.google.ca/books?hl=en&lr=&id=mEBe9MBOWn4C&oi=fnd&pg=PT5&dq=Eccles+%26+Krzus,+2010&ots=KzfhQnMv9n&sig=ECBQCW64JxBHC5OcOvAbyC1FFZM

Eccles, R. G., Krzus, M. P., Rogers, J., & Serafeim, G. (2012). The Need for Sector-Specific Materiality and Sustainability Reporting Standards. Journal of Applied Corporate Finance, 24(2), 65–71. https://doi.org/10.1111/j.1745-6622.2012.00380.x

Eccles, R. G., Serafeim, G., & Krzus, M. P. (2011). Market Interest in Nonfinancial Information. Journal of Applied Corporate Finance, 23(4), 113–127. https://doi.org/10.1111/J.1745-6622.2011.00357.X

Ehrhardt, A., & Nguyen, M. T. (2021). Automated ESG Report Analysis by Joint Entity and Relation Extraction. Communications in Computer and Information Science, 1525 CCIS, 325–340. https://doi.org/10.1007/978-3-030-93733-1_23/FIGURES/11

ESG Seneca, S. (2021, April 22). Sustainability Reporting: GRI and SASB Complement Each Other. Seneca ESG. https://senecaesg.com/insights/sustainability-reporting-gri-and-sasb-complement-each-other/

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd, 96(34), 226–231. https://cdn.aaai.org/KDD/1996/KDD96-037.pdf?source=post_page----------

Fabrizio, K. R., & Kim, E.-H. (2019). Reluctant disclosure and transparency: Evidence from environmental disclosures. Organization Science, 30(6), 1207–1231.

Fasan, M., & Mio, C. (2017). Fostering Stakeholder Engagement: The Role of Materiality Disclosure in Integrated Reporting. Business Strategy and the Environment, 26(3), 288–305. https://doi.org/10.1002/bse.1917

Feldman, R., & Sanger, J. (2007). The text mining handbook: Advanced approaches in analyzing unstructured data. Cambridge university press.

Feng, S., & Gao, L. S. (2020). The Verbal Tone in Mandatory Environmental Disclosures: Evidence from Changes in Disclosures Following SEC Guidance. Social and Environmental Accountability Journal, 40(2), 116–139. https://doi.org/10.1080/0969160X.2020.1719172

Fiandrino, S., & Tonelli, A. (2021). A Text-Mining Analysis on the Review of the Non-Financial Reporting Directive: Bringing Value Creation for Stakeholders into Accounting. Sustainability 2021, Vol. 13, Page 763, 13(2), 763. https://doi.org/10.3390/SU13020763

Fiandrino, S., Tonelli, A., & Devalle, A. (2022). Sustainability materiality research: A systematic literature review of methods, theories and academic themes. Qualitative Research in Accounting & Management, 19(5), 665–695. https://doi.org/10.1108/QRAM-07-2021-0141

Gamsjäger, L., & Ray, D. (2024, June 21). GRI, SASB, CDP – Making sense of overlapping sustainability and climate disclosures. Nordea. https://www.nordea.com/en/news/gri-sasb-cdp-making-sense-of-overlapping-sustainability-and-climate-disclosures

Garst, J., Maas, K., & Suijs, J. (2022). Materiality Assessment Is an Art, Not a Science: Selecting ESG Topics for Sustainability Reports. California Management Review, 65(1), 64–90. https://doi.org/10.1177/00081256221120692

George, G., Osinga, E., Lavie, D., & Scott, B. (2016). Big Data and Data Science Methods for Management Research. Academy of Management Journal, 59(5), 1493–1507. https://doi.org/10.5465/AMJ.2016.4005

Gerber, R., Smit, A., & Botha, M. (2024). An evaluation of environmental, social, and governance reporting in the agricultural sector. Business Strategy & Development, 7(1), e316. https://doi.org/10.1002/bsd2.316

Gerring, J. (2012). Social science methodology: A unified framework (2nd ed). Cambridge University Press.

Goloshchapova, I., Poon, S.-H., Pritchard, M., & Reed, P. (2019). Corporate social responsibility reports: Topic analysis and big data approach. The European Journal of Finance, 25(17), 1637–1654. https://doi.org/10.1080/1351847X.2019.1572637

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press. https://books.google.ca/books?hl=en&lr=&id=omivDQAAQBAJ&oi=fnd&pg=PR5&dq=Goodfellow,+I.,+et+al.+(2016).+%22Deep+Learning.%22+MIT+Press.&ots=MON6insFVW&sig=PvG2nhc8sCFEOrdRO-0KDG9nuOs

Goswami, K., Islam, M. K. S., & Evers, W. (2023). A Case Study on the Blended Reporting Phenomenon: A Comparative Analysis of Voluntary Reporting Frameworks and Standards—GRI, IR, SASB, and CDP. The International Journal of Sustainability Policy and Practice, 19(2), 35.

GRI. (2024). GRI. https://www.globalreporting.org/

GRI and SASB. (2021, April 4). A Practical Guide to Sustainability Reporting Using GRI and SASB Standards. SASB. https://sasb.ifrs.org/knowledge-hub/practical-guide-to-sustainability-reporting-using-gri-and-sasb-standards/

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv:2203.05794. https://doi.org/10.48550/arXiv.2203.05794

Gutierrez-Bustamante, M., & Espinosa-Leal. (2022). Natural Language Processing Methods for Scoring Sustainability Reports—A Study of Nordic Listed Companies. Sustainability, MDPI, 14(15), 1–26.

Hadro, D., Fijałkowska, J., Daszyńska-Żygadło, K., Zumente, I., & Mjakuškina, S. (2022). What do stakeholders in the construction industry look for in non-financial disclosure and what do they get? Meditari Accountancy Research, 30(3), 762–785. https://doi.org/10.1108/MEDAR-11-2020-1093/FULL/XML

Hahn, R., & Kühnen, M. (2013). Determinants of sustainability reporting: A review of results, trends, theory, and opportunities in an expanding field of research. Journal of Cleaner Production, 59, 5–21. https://doi.org/10.1016/j.jclepro.2013.07.005

Hajaya, K. A. (2023). Adherence level to materiality in stand-alone GRI sustainability reporting in Arab nations: Does listing status matter. Afro-Asian J. of Finance and Accounting, 13(3), 369–398. https://doi.org/10.1504/AAJFA.2023.132205

Hamza, S., & Jarboui, A. (2022). CSR or social impression management? Tone management in CSR reports. Journal of Financial Reporting & Accounting, 20(3/4), 599–617. https://doi.org/10.1108/JFRA-04-2020-0115

Harjoto, M. A., Laksmana, I., & Lee, W. E. (2020). Female leadership in corporate social responsibility reporting: Effects on writing, readability and future social performance. Advances in Accounting, 49, 100475. https://doi.org/10.1016/J.ADIAC.2020.100475

Harymawan, I., Ardianto, Ningtyas Soeprajitno, R. R. W., Ratri, M. C., & Paramitasari, Y. I. (2020). Text mining on sustainability reporting: A case study. Journal of Security and Sustainability Issues, 9(May), 48–55. https://doi.org/10.9770/JSSI.2020.9.M(4)

Henriques, R., Gaio, C., & Costa, M. (2022). Sustainability Reporting Quality and Stakeholder Engagement Assessment: The Case of the Paper Sector at the Iberian Level. Sustainability (Switzerland), 14(21), 1–14. https://doi.org/10.3390/su142114404

Ho, J. C., Chen, T. H., & Wu, J. J. (2022). Are corporate social responsibility reports informative? Evidence from textual analysis of banks in China. China Finance Review International, 12(1), 101–120. https://doi.org/10.1108/CFRI-04-2021-0081/FULL/XML

Hoepner, A. G. F., Dimatteo, S., Schaul, J., Yu, P.-S., & Musolesi, M. (2017). Tweeting about Sustainability: Can Emotional Nowcasting Discourage Greenwashing? (SSRN Scholarly Paper 2924088). https://doi.org/10.2139/ssrn.2924088

Hummel, K., & Jobst, D. (2024). An Overview of Corporate Sustainability Reporting Legislation in the European Union. Accounting in Europe, 0(0), 1–36. https://doi.org/10.1080/17449480.2024.2312145

Hummel, K., Mittelbach-Hörmanseder, S., Cho, C. H., & Matten, D. (2022). Corporate social responsibility disclosure: A topic-based approach. Accounting and Business Research. https://doi.org/10.1080/00014788.2022.2071199

Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. Proceedings of the International AAAI Conference on Web and Social Media, 8(1), 216–225. https://ojs.aaai.org/index.php/icwsm/article/view/14550

IFC. (2024). Understanding the Global Reporting Frameworks. International Finance Corporation (IFC), World Bank Group. https://www.ifcbeyondthebalancesheet.org/understanding-global-reporting-frameworks

Ignatov, K. (2021). Sentimental Sustainability: Does What Companies Say Tell More Than What Companies Do? Financial Markets, Institutions & Instruments. https://papers.ssrn.com/abstract=3933185

Ioannou, I., & Serafeim, G. (2017). The consequences of mandatory corporate sustainability reporting. Harvard Business School Research Working Paper, 11–100. https://papers.ssrn.com/SOL3/PAPERS.CFM?ABSTRACT_ID=1799589

Jackson, G., Bartosch, J., Avetisyan, E., Kinderman, D., & Knudsen, J. S. (2020). Mandatory Non-financial Disclosure and Its Influence on CSR: An International Comparison. Journal of Business Ethics, 162(2), 323–342. https://doi.org/10.1007/s10551-019-04200-0

Jacobs, B. W., & Singhal, V. R. (2017). The effect of the Rana Plaza disaster on shareholder wealth of retailers: Implications for sourcing strategies and supply chain governance. Journal of Operations Management, 49–51(1), 52–66. https://doi.org/10.1016/j.jom.2017.01.002

Jafari, S. Q., Shokouhyar, S., & Shokoohyar, S. (2022). Producer-consumer sustainability continuum: Mutual understanding to implement extended producer responsibility. Journal of Cleaner Production, 374, 133880. https://doi.org/10.1016/J.JCLEPRO.2022.133880

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). An Introduction to Statistical Learning: With Applications in R. Springer US. https://doi.org/10.1007/978-1-0716-1418-1

Jaworska, S., & Nanda, A. (2018). Doing Well by Talking Good: A Topic Modelling-Assisted Discourse Study of Corporate Social Responsibility. Applied Linguistics, 39(3), 373–399. https://doi.org/10.1093/APPLIN/AMW014

Johnson, S. C. (1967). Hierarchical clustering schemes. Psychometrika, 32(3), 241–254.

Jørgensen, S., Mjøs, A., & Pedersen, L. J. T. (2022). Sustainability reporting and approaches to materiality: Tensions and potential resolutions. Sustainability Accounting, Management and Policy Journal, 13(2), 341–361. https://doi.org/10.1108/SAMPJ-01-2021-0009

Jurafsky, D. (2000). Speech and Language Processing. Pearson Education India. https://www.thuvienso.net/upload/library/204335/preview/ebook-speech-and-language-processing-an-introduction-to-natural-language-processing-computational-linguistics-and-speech-recognition-part-1-daniel-jur.pdf

Kalra, V., & Aggarwal, R. (2017). Importance of Text Data Preprocessing & Implementation in RapidMiner. ICITKM, 14, 71–75.

Kang, E., & Lam, N. B. (2023). The impact of environmental disclosure on initial public offering underpricing: Sustainable development in Singapore. Corporate Social Responsibility and Environmental Management, 30(1), 119–133. https://doi.org/10.1002/csr.2342

Kang, H., & Kim, J. (2022). Analyzing and visualizing text information in corporate sustainability reports using natural language processing methods. Applied Sciences, 12(11), 5614.

Kang, S. (2021). K-Nearest Neighbor Learning with Graph Neural Networks. Mathematics, 9(8), 830. https://doi.org/10.3390/MATH9080830

Kang, S., & Kim, S. (2022). Lessons Learned from Topic Modeling Analysis of COVID-19 News to Enrich Statistics Education in Korea. Sustainability (Switzerland), 14(6). Scopus. https://doi.org/10.3390/su14063240

Karnani, A. (2010). The case against corporate social responsibility. Wall Street Journal, 23(14), 1–5.

Kayser, S. (2016). The rise and role of the 'Accord' and the 'Alliance' in response to the collapse of Rana Plaza: Real or symbolic change? ERB. https://erb.umich.edu/2016/04/18/the-rise-and-role-of-the-accord-and-the-alliance-in-response-to-the-collapse-of-rana-plaza-real-or-symbolic-change/

Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. The Accounting Review, 91(6), 1697–1724.

Kim, E. H., & Lyon, T. P. (2015). Greenwash vs. Brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. Organization Science, 26(3), 705–723. https://doi.org/10.1287/ORSC.2014.0949/ASSET/IMAGES/LARGE/ORSC.2014.0949-F01.JPEG

Kiriu, T., & Nozaki, M. (2020). A Text Mining Model to Evaluate Firms' ESG Activities: An Application for Japanese Firms. Asia-Pacific Financial Markets, 27(4), 621–632. https://doi.org/10.1007/s10690-020-09309-1

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—A systematic literature review. Information and Software Technology, 51(1), 7–15. https://doi.org/10.1016/j.infsof.2008.09.009

Klimczak, K. M., Hadro, D., & Meyer, M. (2023). Executive communication with stakeholders on sustainability: The case of Poland. Accounting in Europe. https://doi.org/10.1080/17449480.2023.2213242

Knight, W. (2021). The Left and the Right Speak Different Languages—Literally. https://www.wired.com/story/left-right-speak-different-languages-literally/

Kountouri, I., Manousakis, E., & Tsekrekos, A. E. (2019). Latent semantic analysis of corporate social responsibility reports (with an application to Hellenic firms). International Journal of Disclosure and Governance, 16(1), 1–19. https://doi.org/10.1057/s41310-018-0053-z

Kozlowski, A., Searcy, C., & Bardecki, M. (2015). Corporate sustainability reporting in the apparel industry an analysis of indicators disclosed. International Journal of Productivity and Performance Management, 64(3), 377–397. https://doi.org/10.1108/IJPPM-10-2014-0152/FULL/XML

KPMG international. (2014). The essentials of materiality assessment—KPMG China. KPMG. https://home.kpmg/cn/en/home/insights/2014/10/materiality-assessment.html

KPMG International. (2022). Big shifts, small steps> Survey of Sustainability Reporting 2022. https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2023/04/big-shifts-small-steps.pdf

Krippendorff, K. (2019). Content Analysis: An Introduction to Its Methodology. SAGE Publications, Inc. https://doi.org/10.4135/9781071878781

Kvasničková Stanislavská, L., Pilař, L., Fridrich, M., Kvasnička, R., Pilařová, L., Afsar, B., & Gorton, M. (2023). Sustainability reports: Differences between developing and developed countries. Frontiers in Environmental Science, 11. https://doi.org/10.3389/FENVS.2023.1085936/FULL

Labowitz, S., & Baumann-Pauly, D. (2014). Business as usual is not an option. New York University Center for Business and Human Rights. https://www.stern.nyu.edu/sites/default/files/assets/documents/con_047408.pdf

Lai, A., & Stacchezzini, R. (2021). Organisational and professional challenges amid the evolution of sustainability reporting: A theoretical framework and an agenda for future research. Meditari Accountancy Research, 29(3), 405–429. https://doi.org/10.1108/MEDAR-02-2021-1199

Laufer, W. S. (2003). Social Accountability and Corporate Greenwashing. Journal of Business Ethics, 43(3), 253–261. https://doi.org/10.1023/A:1022962719299

Leavitt, K., Schabram, K., Hariharan, P., & Barnes, C. M. (2021). Ghost in the machine: On organizational theory in the age of machine learning. Academy of Management Review, 46(4), 750–777.

Lewis, C., & Young, S. (2019). Fad or future? Automated analysis of financial text and its implications for corporate reporting. Accounting and Business Research, 49(5), 587–615. https://doi.org/10.1080/00014788.2019.1611730

Li, C., & Huang, M. (2023). Environmental Science and Pollution Research. https://doi.org/10.1007/s11356-023-30301-5

Li, M., & Zhao, L. (2021). Exploring Global Fashion Sustainability Practices through Dictionary-Based Text Mining. Clothing and Apparels Research Journal, 1–16. https://doi.org/10.1177/0887302X21998268

Li, Y., & Rockinger, M. (2024). Unfolding the Transitions in Sustainability Reporting. Sustainability, 16(2), Article 2. https://doi.org/10.3390/su16020809

Liang, X., & Wu, H. (2022). Does the Tone in Corporate Social Responsibility Reports Misdirect Analysts' Forecasts in China? Sustainability, 14(24), 16631. https://doi.org/10.3390/SU142416631

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R News, 2(3), 18–22.

Liew, W. T., Adhitya, A., & Srinivasan, R. (2014). Sustainability trends in the process industries: A text mining-based analysis. Computers in Industry, 65(3), 393–400. https://doi.org/10.1016/j.compind.2014.01.004

Lindgren, C., Huq, A. M., & Carling, K. (2021a). Who Are the Intended Users of CSR Reports? Insights from a Data-Driven Approach. Sustainability, 13(3), 1070. https://doi.org/10.3390/su13031070

Liu, J., Sun, J., & Fang, M. (2022). Construction and Analysis of Corporate Social Responsibility Report and Equity Capital Cost Algorithm Model Based on Intelligent Computing. Security and Communication Networks, 2022. https://doi.org/10.1155/2022/3306735

Liu, S.-H., Chen, S.-Y., & Li, S.-T. (2017). Text-Mining Application on CSR Report Analytics: A Study of Petrochemical Industry. 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), 76–81. https://doi.org/10.1109/IIAI-AAI.2017.164

Locke, R., Amengual, M., & Mangla, A. (2009). Virtue out of Necessity? Compliance, Commitment, and the Improvement of Labor Conditions in Global Supply Chains. Politics & Society, 37(3), 319–351. https://doi.org/10.1177/0032329209338922

Lu, J., & Jagoda, K. (2023). Sentiment analysis in sustainability accounting reporting: Does the tone reveal future environmental performance? International Journal of Accounting and Finance. https://www.inderscienceonline.com/doi/10.1504/IJAF.2023.134518

Luo, X., Wang, H., Raithel, S., & Zheng, Q. (2015). Corporate social performance, analyst stock recommendations, and firm future returns. Strategic Management Journal, 36(1), 123–136. https://doi.org/10.1002/SMJ.2219

Lyon, T. P., & Maxwell, J. W. (2011). Greenwash: Corporate Environmental Disclosure under Threat of Audit. Journal of Economics & Management Strategy, 20(1), 3–41. https://doi.org/10.1111/J.1530-9134.2010.00282.X

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1(14), 281–297.

Mangsor, N. S. M. N., Nasir, S. A. M., Yaacob, W. F. W., Ismail, Z., & Rahman, S. A. (2022). Analysing corporate social responsibility reports using document clustering and topic modeling techniques. Indonesian Journal of Electrical Engineering and Computer Science, 26(3), 1546–1555. https://doi.org/10.11591/IJEECS.V26.I3.PP1546-1555

Mani, V., Gunasekaran, A., & Delgado, C. (2018). Supply chain social sustainability: Standard adoption practices in Portuguese manufacturing firms. International Journal of Production Economics, 198, 149–164. https://doi.org/10.1016/J.IJPE.2018.01.032

Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to Information Retrieval. Cambridge University Press.

Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, Norms, and Selective Disclosure: A Global Study of Greenwashing. Organization Science, 27(2), 483–504. https://doi.org/10.1287/orsc.2015.1039

Maulida, S., & Hakim, B. A. (2022). Twitter Sentiment Analysis on Green Finance. Accounting and Sustainability, 1(1), Article 1. https://doi.org/10.58968/as.v1i1.256

Merkl-Davies, D. M., & Brennan, N. M. (2008). Discretionary Disclosure Strategies in Corporate Narratives: Incremental Information or Impression Management? (SSRN Scholarly Paper 1089447). https://papers.ssrn.com/abstract=1089447

Michelon, G., Pilonato, S., & Ricceri, F. (2015). CSR reporting practices and the quality of disclosure: An empirical analysis. Critical Perspectives on Accounting, 33, 59–78. https://doi.org/10.1016/j.cpa.2014.10.003

Michelon, G., Trojanowski, G., & Sealy, R. (2022). Narrative Reporting: State of the Art and Future Challenges. Accounting in Europe, 19(1), 7–47. https://doi.org/10.1080/17449480.2021.1900582

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space (arXiv:1301.3781). arXiv. http://arxiv.org/abs/1301.3781

Miller, H. G., & Mork, P. (2013). From data to decisions: A value chain for big data. It Professional, 15(1), 57–59.

Miner, G. D., Elder, J., Fast, A., Hill, T., Nisbet, R., & Delen, D. (2012). Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications. Academic Press.

Modapothala, J. R., & Issac, B. (2009). Angin. Software Technology and Engineering, 151–155. https://doi.org/10.1142/9789814289986_0028

Mohan, V., Vijayarani, S., & Janani, M. R. (2016). Advanced Computational Intelligence. An International Journal (ACII), 3(1). https://doi.org/10.5121/acii.2016.3104

Moreno, A. I., & Caminero, T. (2022). Application of text mining to the analysis of climate-related disclosures. International Review of Financial Analysis, 83, 102307. https://doi.org/10.1016/J.IRFA.2022.102307

Mućko, P. (2021). Sentiment analysis of CSR disclosures in annual reports of EU companies. Procedia Computer Science, 192, 3351–3359. https://doi.org/10.1016/j.procs.2021.09.108

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. Journal of the American Medical Informatics Association : JAMIA, 18(5), 544–551. https://doi.org/10.1136/amiajnl-2011-000464

Nast, C. (2023, April 24). 10 years since Rana Plaza, not enough has changed. Vogue Business. https://www.voguebusiness.com/sustainability/10-years-since-rana-plaza-not-enough-has-changed

Ngu, S. B., & Amran, A. (2021). Materiality disclosure in sustainability reporting: Evidence from Malaysia. Asian Journal of Business and Accounting, 14(1), 225–252. https://doi.org/10.22452/ajba.vol14no1.9

Ning, X., Yim, D., & Khuntia, J. (2021). Online Sustainability Reporting and Firm Performance: Lessons Learned from Text Mining. Sustainability, 13(3), 1069. https://doi.org/10.3390/su13031069

Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. International Journal of Information Management, 53, 102104. https://doi.org/10.1016/j.ijinfomgt.2020.102104

Niveditha, R., Ns, N. K., Parimi, M. R., Aaditya Raam, S., & Babu, S. (2020). Develop CSR Themes using Text-Mining and Topic Modelling Techniques. Proceedings - 2020 IEEE International Conference on Cloud Computing in Emerging Markets, CCEM 2020, 67–71. https://doi.org/10.1109/CCEM50674.2020.00024

Ortar, L. (2018). Materiality matrixes in sustainability reporting: An empirical examination. Available at SSRN 3117749. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3117749

Padilla-Garrido, N., Aguado-Correa, F., Rabadán-Martín, I., López-Jiménez, J. M., De La Vega-Jiménez, J. J., & Peletier-Ribera, I. (2024). Materiality analysis in sustainability reporting: Insights from large Spanish companies. Corporate Social Responsibility and Environmental Management, csr.2866. https://doi.org/10.1002/csr.2866

Pan, Y. (2016). Sustainability trends in financial services sector: Evidence from Europe and North America. 2016 13th International Conference on Service Systems and Service Management, ICSSSM 2016. https://doi.org/10.1109/ICSSSM.2016.7538549

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval, 2(1–2), 1–135.

Pedhazur, E. J., & Schmelkin, L. P. (2013). Measurement, Design, and Analysis: An Integrated Approach. Psychology Press.

Picard, Ohl, & Beagent. (2023). Why impact matters for company valuations. PwC. https://www.pwc.com/gx/en/services/audit-assurance/corporate-reporting/esg-reporting/esg-impact-company-valuation.html

Pimonenko, T., Bilan, Y., Horák, J., Starchenko, L., & Gajda, W. (2020). Green Brand of Companies and Greenwashing under Sustainable Development Goals. Sustainability, 12(4), Article 4. https://doi.org/10.3390/su12041679

Pizzi, S., Principale, S., & de Nuccio, E. (2022). Material sustainability information and reporting standards. Exploring the differences between GRI and SASB. Meditari Accountancy Research, 31(6), 1654–1674. https://doi.org/10.1108/MEDAR-11-2021-1486

Pollach, I. (2018). Issue Cycles in Corporate Sustainability Reporting: A Longitudinal. Environmental Communication, 12(2), 247–260. https://doi.org/10.1080/17524032.2016.1205645

Raghupathi, V., Ren, J., & Raghupathi, W. (2020). Identifying Corporate Sustainability Issues by Analyzing Shareholder Resolutions: A Machine-Learning Text Analytics Approach. Sustainability, MDPI, 12(11), 4753. https://doi.org/10.3390/SU12114753

Ramakrishnan, J., Liu, T., Zhang, F., Seshadri, K., Yu, R., & Gou, Z. (2023a). A decision tree-based modeling approach for evaluating the green performance of airport buildings. Environmental Impact Assessment Review, 100, 107070. https://doi.org/10.1016/J.EIAR.2023.107070

Raman, N., Bang, G., & Nourbakhsh, A. (2020). Mapping ESG trends by distant supervision of neural language models. Machine Learning and Knowledge Extraction, 4, 453–468. https://doi.org/10.3390/make2040025

Reinecke, J., & Donaghey, J. (2015). After Rana Plaza: Building coalitional power for labour rights between unions and (consumption-based) social movement organisations. Organization (London, England), 22(5), 720–740. https://doi.org/10.1177/1350508415585028

Resnik, S., & Koklič, M. K. (2018). User-generated tweets about global green brands: A sentiment analysis approach. Trziste= Market, 30(2), 125–145.

Reuter, N., Debortoli, S., & Müller, O. (2014). Identifying The Role Of Energy-Related Environmental Sustainability Using Text Mining. Ecis, 0–15.

Rhoden, I., Ball, C. S., Vögele, S., & Kuckshinrichs, W. (2023). Minding the gap-relating disclosure to contexts of sustainability reporting in the automotive industry. Corporate Social Responsibility and Environmental Management, 30(2), 846–857. https://doi.org/10.1002/CSR.2392

Riani, R., & Rusydiana, A. S. (2023). Exploring Sentiment Analysis of Sustainable Finance Initiatives: A Text Mining Approach. Accounting and Sustainability, 2(1), Article 1. https://doi.org/10.58968/as.v2i1.266

Ribeiro, F. N., Araújo, M., Gonçalves, P., André Gonçalves, M., & Benevenuto, F. (2016). SentiBench—A benchmark comparison of state-of-the-practice sentiment analysis methods. EPJ Data Science, 5(1), 23. https://doi.org/10.1140/epjds/s13688-016-0085-1

Roca, L. C., & Searcy, C. (2012). An analysis of indicators disclosed in corporate sustainability reports. Journal of Cleaner Production, 20(1), 103–118. https://doi.org/10.1016/J.JCLEPRO.2011.08.002

Rocca, L., Giacomini, D., & Zola, P. (2020). Environmental disclosure and sentiment analysis: State of the art and opportunities for public-sector organisations. Meditari Accountancy Research, 29(3), 617–646. https://doi.org/10.1108/MEDAR-09-2019-0563/FULL/HTML

Saha, A., & Nabareseh, S. (2015). Communicating corporate social responsibilities: Using text mining for a comparative analysis of banks in India and Ghana. Mediterranean Journal of Social Sciences, 6(3), 11–20. https://doi.org/10.5901/mjss.2015.v6n3s1p11

Sai, P. K., Gupta, P., & Fernandes, S. F. (2019). Analysing Performance of Company through Annual reports using Text Analytics. 21–31. https://doi.org/10.1109/ICD47981.2019.9105673

Sambhanthan, A., & Potdar, V. (2015). Green Business Practices for Software Development Companies: An Explorative Text Analysis of Business Sustainability Reports. International Journal of Enterprise Information Systems, 11(3), 13–26. https://doi.org/10.4018/IJEIS.2015070102

Sambhanthan, A., & Potdar, V. (2016). Waste management strategies for Software Development companies: An explorative text analysis of business sustainability reports. 2016 IEEE 14th International Conference on Software Engineering Research, Management and Applications (SERA), 179–184. https://doi.org/10.1109/SERA.2016.7516144

Samkin, G. (2012). Changes in sustainability reporting by an African defence contractor: A longitudinal analysis. Meditari Accountancy Research, 20(2), 134–166. https://doi.org/10.1108/10222521211277834/FULL/XML

SASB. (2021). SASB. https://sasb.org/

SASB. (2024). SASB. https://sasb.ifrs.org/

Shahi, A. M., Issac, B., & Modapothala, J. R. (2012). Enhanced intelligent text categorization using concise keyword analysis. ICIMTR 2012 - 2012 International Conference on Innovation, Management and Technology Research, 574–579. https://doi.org/10.1109/ICIMTR.2012.6236461

Shahi, A. M., Issac, B., & Modapothala, J. R. (2014). Automatic analysis of corporate sustainability reports and intelligent scoring. International Journal of Computational Intelligence and Applications, 13(1). https://doi.org/10.1142/S1469026814500060

Sharma, A., Kaushik, K., Awasthy, P., & Gawande, A. (2022). Leveraging Text Mining for Trend Analysis and Comparison of Sustainability Reports: Evidence from Fortune 500 Companies. American Business Review, 25(2), 8. https://doi.org/10.37625/abr.25.2.416-438

Slacik, J., & Greiling, D. (2020). Compliance with materiality in G4-sustainability reports by electric utilities. International Journal of Energy Sector Management, 14(3), 583–608.

Smeuninx, N., De Clerck, B., & Aerts, W. (2016). Measuring the Readability of Sustainability Reports: A Corpus-Based Analysis Through Standard Formulae and NLP. International Journal of Business Communication, 57(1), 52–58. https://doi.org/10.1177/2329488416675456

Son, B.-G. (2005). WordStat 5.0: Content analysis software offers operations researchers valuable tool for exploring vast, untapped amounts of textual data. OR/MS Today, 32(5), 54–59.

Stanik, C., Pietz, T., & Maalej, W. (2021). Unsupervised Topic Discovery in User Comments. 2021 IEEE 29th International Requirements Engineering Conference (RE), 150–161. https://doi.org/10.1109/RE51729.2021.00021

Suddaby, R., & Panwar, R. (2022). On the Complexity of Managing Transparency. California Management Review, 65(1), 5–18. https://doi.org/10.1177/00081256221128766

Sustainability Knowledge Group. (2021). GRI and SASB Reporting Standards: Working Together. Sustainability Knowledge Group. https://sustainabilityknowledgegroup.com/gri-and-sasb-reporting-standards-working-together/

Szekely, N., & Vom Brocke, J. (2017). What can we learn from corporate sustainability reporting? Deriving propositions for research and practice from over 9,500 corporate sustainability reports published between 1999 and 2015 using topic modelling technique. PLOS ONE, 12(4), e0174807. https://doi.org/10.1371/JOURNAL.PONE.0174807

Tate, W., Ellram, L., & Kirchoff, J. (2009). Corporate social responsibility reports: A thematic analysis related to supply chain management. Journal of Supply Chain Management, 46(1), 19–44. https://doi.org/10.1111/j.1745-493X.2009.03184.x

Thangam, D., Chavadi, C., & Manoharan, S. (n.d.). Corporate Sustainability, Greenwash, and Its Impact on Environment, Society, and Governance. Retrieved August 30, 2024, from https://www.researchgate.net/publication/381584492_CORPORATE_SUSTAINABILITY_GREENWASH_AND_ITS_IMPACT_ON_ENVIRONMENT_SOCIETY_AND_GOVERNANCE_CSG-ESG_2024

Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. Journal of the American Society for Information Science and Technology, 63(1), 163–173. https://doi.org/10.1002/asi.21662

Thelwell, K. (2021, May 26). The Impact of Fast Fashion in Bangladesh. The Borgen Project. https://borgenproject.org/fast-fashion-in-bangladesh/

Thomas, D. A. (2014). Searching for Significance in Unstructured Data: Text Mining with Leximancer. European Educational Research Journal, 13(2), 235–256. https://doi.org/10.2304/eerj.2014.13.2.235

Tiscini, R., Martiniello, L., & Lombardi, R. (2021). Circular economy and environmental disclosure in sustainability reports: Empirical evidence in cosmetic companies. Business Strategy and Environment, 31(3), 892–907. https://doi.org/10.1002/bse.2924

Tomlinson, B., & Godshall, L. (2023). ESG regulation gives materiality a bigger stake. EY. https://www.ey.com/en_us/insights/assurance/esg-regulation-gives-materiality-a-bigger-stake

Torelli, R., Balluchi, F., & Lazzini, A. (2020). Greenwashing and environmental communication: Effects on stakeholders' perceptions. Business Strategy and the Environment, 29(2), 407–421. https://doi.org/10.1002/bse.2373

Tóth, Á., & Suta, A. (2021). Global sustainability reporting in the automotive industry via the eXtensible business reporting language. Chemical Engineering Transactions, 88, 1087–1092. https://doi.org/10.3303/CET2188181

Tóth, Á., Suta, A., & Szauter, F. (2022). Interrelation between the climate-related sustainability and the financial reporting disclosures of the European automotive industry. Clean Technologies and Environmental Policy, 24(1), 437–445. https://doi.org/10.1007/S10098-021-02108-W/TABLES/4

Tóth, Á., Szigeti, C., & Suta, A. (2021). Carbon Accounting Measurement with Digital Non-Financial Corporate Reporting and a Comparison to European Automotive Companies Statements. Energies, 14(18), 5607. https://doi.org/10.3390/en14185607

Tremblay, M. C., Mason, R. A., & Forsgren, N. (2021). Theories in Flux: Reimagining Theory Building in the Age of Machine Learning. MIS Quarterly, 45(1), 455–459. https://doi.org/10.25300/MISQ/2021/15434.1.1

Tremblay, M. C., Parra, C. M., & Castellanos, A. (2015). Corporate social responsibility reports: Understanding topics via text mining. 2015 Americas Conference on Information Systems, AMCIS 2015.

Uyar, A., Koseoglu, M. A., Kılıç, M., & Mehraliyev, F. (2021). Thematic structure of sustainability reports of the hospitality and tourism sector: A periodical, regional, and format-based analysis. Current Issues in Tourism, 24(18), 2602–2627. https://doi.org/10.1080/13683500.2020.1847050

Vaio, A. D., Hassan, R., D'Amore, G., & Strologo, A. D. (2022). Digital Technologies for Sustainable Waste Management On-Board Ships: An Analysis of Best Practices From the Cruise Industry. IEEE Transactions on Engineering Management. https://doi.org/10.1109/TEM.2022.3197241

Vieira, I. L., Silva, E. R. da, Martini, L. C. de, & Rangel, L. A. D. (2021). Proposal for an analytical model of materiality in the Global Reporting Initiative Standards reports. Production, 31, e20210003.

Vigneau, L., & Adams, C. A. (2023). The failure of transparency as self-regulation. Sustainability Accounting, Management and Policy Journal, 14(4), 852–876.

Vouros, P., Nomikos, S., Halkos, G., Evangelinos, K., Sfakianaki, E., Konstandakopoulou, F., Fotiadis, S., Karagiannis, I., Skouloudis, A., & Nikolaou, I. E. (2020). Introducing fundamental accountability principles in sustainability reporting assessment: A cross-sectoral analysis from the Greek business sector. Environmental Quality Management, 29(4), 33–50. https://doi.org/10.1002/tqem.21689

Wallbaum, H., Krank, S., & Teloh, R. (2011). Prioritizing Sustainability Criteria in Urban Planning Processes: Methodology Application. Journal of Urban Planning and Development, 137(1), 20–28. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000038

Wang, W., Kang, W., & Mu, J. (2023). Mapping research to the sustainable development goals (sdgs). https://www.researchsquare.com/article/rs-2544385/latest

Wang, X., Yuen, K. F., Wong, Y. D., & Li, K. X. (2020). How can the maritime industry meet Sustainable Development Goals? An analysis of sustainability reports from the social entrepreneurship perspective. Transportation Research Part D: Transport and Environment, 78, 102173. https://doi.org/10.1016/J.TRD.2019.11.002

Wanner, J., & Janiesch, C. (2019). Big data analytics in sustainability reports: An analysis based on the perceived credibility of corporate published information. Business Research, 12(1), 143–173. https://doi.org/10.1007/s40685-019-0088-4

Whittingham, K. L., Earle, A. G., Argiolas, A., & Hiz, D. I. L. la. (2022). The impact of the United Nations Sustainable Development Goals on corporate sustainability reporting. BRQ Business Research Quarterly, 26(1), 45–61. https://doi.org/10.1177/23409444221085585

Wu, P.-J., & Huang, P.-C. (2018). Business analytics for systematically investigating sustainable food supply chains. Journal of Cleaner Production, 203, 968–976.

Wu, S. R., Shao, C., & Chen, J. (2018). Approaches on the Screening Methods for Materiality in Sustainability Reporting. Sustainability, MDPI, 10(9), 3233. https://doi.org/10.3390/SU10093233

Yang, B. M., & Yang, O. S. (2022). Assessing the Effect of Dynamic Capabilities on the ESG Reporting and Corporate Performance Relationship With Topic Modeling: Evidence From Global Companies. Frontiers in Psychology, 13, 898935. https://doi.org/10.3389/FPSYG.2022.898935/BIBTEX

Ye, M., Lu, W., Flanagan, R., & Chau, K. W. (2019). Corporate social responsibility "glocalisation": Evidence from the international construction business. Corporate Social Responsibility and Environmental Management, 27(2), 655–669. https://doi.org/10.1002/csr.1831

Ye, M., Lu, W., & Xue, F. (2022). Impact of Institutional Distance on Environmental and Social Practices in Host Countries: Evidence from International Construction Companies. Journal of Construction Engineering and Management, 148(1). https://doi.org/10.1061/(ASCE)CO.1943-7862.0002226

Yu, E. P., Luu, B. V., & Chen, C. H. (2020). Greenwashing in environmental, social and governance disclosures. Research in International Business and Finance, 52, 101192. https://doi.org/10.1016/j.ribaf.2020.101192

Zakaria, M., Aoun, C., & Liginlal, D. (2021). Objective sustainability assessment in the digital economy: An information entropy measure of transparency in corporate sustainability reporting. Sustainability, MDPI, 13(3), 1–18. https://doi.org/10.3390/su13031054

Zharfpeykan, R., & Akroyd, C. (2023). Evaluating the outcome effectiveness of the global reporting initiative transitions. Sustainability Accounting, Management and Policy Journal, 14(6), 1101–1125.

Zhou, A. (2021). Communicating corporate LGBTQ advocacy: A computational comparison of the global CSR discourse. Public Relations Review, 47(4), 102061. https://doi.org/10.1016/J.PUBREV.2021.102061

Zhou, K., Jin, X., Li, X., & Tao, Y. (2023). Enhancing sustainable development through effective disclosure: Corporate environmental performance and readability. Corporate Social Responsibility and Environmental Management. https://doi.org/10.1002/CSR.2569

Zhou, Y., Li, X., Wang, X., & Yuen, K. F. (2022). Intelligent container shipping sustainability disclosure via stakeholder sentiment views on social media. Marine Policy, 135. https://doi.org/10.1016/j.marpol.2021.104853

Zhou, Y., Wang, X., & Yuen, K. F. (2021). Sustainability disclosure for container shipping: A text-mining approach. Transport Policy, 110, 465–477. https://doi.org/10.1016/j.tranpol.2021.06.020

Conclusion

This thesis set out to explore the transformative potential of automated text analysis techniques, particularly text mining (TM) and natural language processing (NLP), in the context of sustainability reporting (SR). By focusing on two interrelated studies, this research contributes significantly to both the academic understanding and practical application of these advanced technologies in enhancing the analysis and communication of sustainability data.

The first study presented a systematic literature review (SLR) of TM and NLP techniques applied to SRs. The review provided a thorough assessment of the existing methodologies, research objectives, and the extent to which these automated techniques have been utilized in the field. The findings revealed that while there has been progress in adopting these technologies, the full potential of TM-NLP to extract meaningful insights from the increasing volume of SR data remains largely untapped. The study underscored the need for further refinement in both the depth and breadth of analysis, suggesting that future research should focus on overcoming the current limitations to fully realize the benefits of automated text analysis in SRs.

Building on the insights from the literature review, the second study applied advanced text mining techniques to a specific case study within the apparel industry, particularly focusing on companies implicated in the Rana Plaza disaster. This study demonstrated the practical application of TM-NLP in assessing materiality and balance within SRs. The findings highlighted a significant bias towards positive reporting, with companies often emphasizing strengths and opportunities while downplaying risks and challenges. This imbalance raises critical concerns about the credibility and transparency of SRs, especially in industries where ethical and social issues are paramount. The study not only advanced methodological approaches but also provided a model for future research in theme extraction and sentiment analysis, offering deeper insights into the authenticity of corporate sustainability claims.

Together, these two studies contribute to a deeper understanding of how automated text analysis can be leveraged to improve the quality and reliability of sustainability reports. The research highlights the dual need for both comprehensive and nuanced analysis to ensure that SRs serve their intended purpose of providing transparent, balanced, and accurate information to stakeholders. By addressing the identified research gaps, this thesis offers practical recommendations for companies, policymakers, and researchers aiming to enhance the effectiveness of sustainability reporting.

In conclusion, this thesis makes a significant contribution to the field of sustainability reporting by demonstrating the potential of TM-NLP techniques in both theoretical and practical contexts. The insights gained from this research pave the way for more robust and reliable SR practices, ultimately fostering greater transparency and accountability in corporate sustainability efforts. As the demand for high-quality sustainability reporting continues to grow, the findings of this thesis underscore the importance of advancing automated text analysis methods to meet the evolving needs of stakeholders and society at large.