

Communication Practices of Corporate Social Responsibility: LLM Informed Discourse Analysis

Rationale:

Sustainability is one of the most pressing and complex challenges of our time. It requires a global perspective and a systemic approach that considers the interrelated economic, social, and environmental dimensions of human well-being. In this context, businesses have a crucial role and responsibility to contribute to sustainable development and to meet the expectations and demands of their stakeholders. As organisations face complex problems, from climate change to social and economic inequality, they need to show real commitment to sustainability - but as incessant corporate scandals attest, that is clearly not the case. From gaslighting (famously the oil industry, Hicks 2022) to greenwashing (including world-brands like Adidas, Mathews 2021), business organisations employ a range of communication strategies to obscure their role in actually causing societal harm (like Facebook, Ables, 2023) or environmental damage (famously Shell, Esteban and McCardle 2021).

The importance of language

Much - if not all - of the information that stakeholders receive about sustainability comes from corporate communication channels, such as corporate social responsibility (CSR) reports, websites, press releases, and social media. These channels are often used by businesses to shape their image and reputation, to influence public opinion and policy, and to justify or conceal their questionable practices. In these contexts, communication has two crucial roles. Communication - language and other meaning making resources - do not just represent aspects of reality but actively create shared meanings and perceptions about the organisations actions and intentions in sustainable development (Putnam, Phillips & Chapman, 1996). On the other hand, communication is instrumental, used to persuade or manipulate stakeholders' perceptions and paint an image of corporate responsibility, sometimes even at the cost of truthfulness.

The importance of focus on communicative practices

The blurred line between genuine intention and action, and misleading representation can often be tracked down if we look closely at how an organisation communicates about CSR. Consider, for example, the Volkswagen Emissions Scandal (Dieselgate). Volkswagen, in its communication campaigns, actively promoted their diesel vehicles as low-emission, environmentally friendly options, leveraging their "Clean Diesel" branding to shape public perceptions. However, in 2015, investigations unveiled that Volkswagen had employed 'defeat devices' to manipulate emissions tests, which lead to a disparity between the company's communicated image and the actual environmental impact of their cars (Painter and Martins 2017). As Painter's and Martin's project (2017), along with other discourse-focussed scholarship in CSR communication (see Catenaccio 2012) demonstrates, the close examination of textual artefacts is an effective way to expose contradictions between imagined and real responsibility. This is the basic realisation that influences this project. The project is also motivated by the supervisory teams' deep

commitment to develop ways that can help us hold organisations accountable and lead to true sustainable practices.

New approaches to discourse analysis of CSR

Considering the importance of language and communication, it is unsurprising that researchers have been looking at different communication artefacts to expose agendas of CSR like websites (Pollach, 2003), CEO communication (Rajandra & Taib, 2014) or organisational CSR reports (Gill & Broderick 2014). Previous studies predominantly adopt qualitative or corpus-assisted discourse analytical approaches, (Aleizza, 2015) or focus on comparisons between e.g. national cultures (Castelo 2014) or industries (O'Connor et al. 2017, Song & Wen 2020). While this work makes crucial contributions to directing our attention to discourse strategies prevalent in genuine and misleading CSR discourse, small scale studies are not always sufficient to explore the nuance of (implied) meaning across a large data set, or to develop models that could highlight potential discourse elements that may predict questionable (communication) practices. This is precisely the gap this proposed project aims to fulfill. In the last few years, there has been a revolution in the field of Natural Language Processing due to the introduction of Large Language Models (LLMs). These types of models are general-purpose, i.e. they are capable of performing a wide range of tasks, including tasks they have never seen before. Moreover, a crucial difference with the previous set of NLP models is that they can receive instructions in natural language. We now have access to models that can analyse large quantities of text automatically, while doing this in a fashion that we specified in detail beforehand. However, given how new they still are, knowing what the best practices are is still an open research question.

Research Questions

To achieve the aims described below, the project sets out to answer the following research questions:

- How are discourse strategies and patterns of communication linked to genuine vs. misleading CSR communication practices?
- How can LLMs be trained and fine-tuned to identify discourse strategies representative of (non)genuine CSR communication?
- How can LLMs be employed to predict the intentions of a CSR communication artefact?
- In what ways do inherent biases of LLMs influence the analysis and how can this be mitigated in the project's methodology?

Aims of the project

The project has 3 interrelated aims:

- To expose the crucial role of language in genuine and misleading sustainability communication,
- To develop and test LLM-informed discourse analytical methods;
- To contribute to the growing knowledge about inherent biases in LLMs and how these may affect data analysis.

Methodology

While discourse-focussed studies in CSR communication have primarily adopted either qualitative or corpus-assisted discourse analytic approaches, these methods may pose limitations in scope and scale. The “old school” method of corpus linguistics is to look for things we explicitly define. This was often based on grammar, sentence structure and syntax. Though this can work well in practice, you generally only find what you put in. The new generation of LLMs might be able to go beyond this. We can still show them the explicit examples, but then instruct them to find other examples that are similar in spirit, but are potentially quite different in terms of syntax or surface realisation. In addition, the changing landscape of corporate communication - including for example and increasing adoption of various digital and multimodal communication channels, is a challenge of not just comprehensiveness but also adaptability of methods. This is where LLMs may offer a fundamental paradigm shift in discourse analysis. There is already emerging literature that experiments with and promotes the use of LLMs to explore linguistic patterns and nuances of language use. Törnberg (2023) for example proposes a step-by step prompt engineering method to train LLMs in quantitative text analysis in social sciences. Fan and Jiang (2023) demonstrate a method of training LLMs to conduct discourse analysis on interactions and show the precision of the annotation the LLM was able to achieve. This emerging work points in the direction that LLMs can be used as tools to assist in discourse analytical work.

However, thus far there is a lacuna in research that would 1) describe the process of the development of prompts based on previously existing discourse analytical methodologies, and 2) provide evidence of the effectiveness of these prompts. Answering these questions will be a unique contribution of this research project both for theoretical understanding and more analytical method development.

We are mindful that LLMs reflect the biases present in their training data. When analysing something as nuanced and potentially manipulative as corporate communication, it is important to consider the possible biases in interpretation. To address this, we have included a sub-topic to specifically address inherent biases in LLMs and how they may affect data analysis. We also have planned mitigation strategies regarding the texts we analyze, as we will discuss below.