1. **Description**

Brands are turning to social media to stimulate customer-mindset, building, and extending their brand positioning. While brands share their core positioning in their mission and vision statements (and through their marketing activities), they further reinforce it or extend it to capture new dimensions through their social media. This allows brands to keep up with the evolution of consumers and the marketplace (e.g., aligning their brands with the UN sustainable development goals (SDG), new GDPR policies, or highlighting ways to face the post-pandemic crisis). In this research, we focus on the relationship between presentation and perception of brand positioning, as previous work has shown that this relationship can affect brand evaluations on the stock exchange. We assess the alignment or misalignment between the brand positioning advocated by brands in their mission statements and their actual social media content that contributes to position the brand in the mind of the consumers. This can help managers to identify areas of misalignment to prevent alienating consumers who may stop engaging with brands that they perceive do not align with their values. Moreover, this is relevant for organizations such as B Lab, who certifies B Corporations (companies that meet high standards of social and environmental performance). Understanding if certified B Corporations not only follow sustainable practices, but also take an active role in “educating” consumers on sustainability via their social media (even just by mentioning it) may provide a new way to track their societal impact.

1. **Scientific project description**

To capture brand positioning, we extracted the mission statements from the websites of 200 leading global brands (including certified B corps) in 2021 and 2022. Next, we retrieved the tweets posted by the brands and by the consumers (about these brands) from 01.01.2020 to 31.12.2022. We analyze the mission statements, brand tweets, and consumer tweets with a text-mining approach to construct count-based scores. We develop dictionaries to capture brand positioning from social media, comparing the importance of the 17 UN SDG dimensions (e.g., community, education, equality, sustainability, well-being) to traditional brand positioning dimensions (e.g., quality, price). We provide a score for every brand and their competitors to verify the alignment of social media content to a brand positioning strategy, as well as the relevance of every dimension over time. Moreover, using a seeded Latent Dirichlet Allocation topic modeling approach, we further identify relevant brand positioning dimensions and validate our theory-driven approach incorporating our dictionaries in the construction and interpretation of the topics. Finally, we link brand positioning in the mind of the consumers (and its alignment with the brand official mission and its social media content) to customer engagement (likes, comments, and shares) and brand followers, identifying different strategies across brands and industries. Thanks to our collaboration with B Lab, we will share our insights with B Lab and their certified B Corporations, hopefully to showcase the value of more sustainable marketing practices.

The Small NWO request will be primarily used for the data analysis of the consumer tweets as we cannot perform these on our university laptops. The consumer tweets do not entail any personal data and will only be used in the analysis on an aggregated level by brand and year. We will not report any personal data in our research and only present our findings on an aggregated level. The research will be based on heavy computing method that are required in top-tier high impact marketing journals such as Journal of Marketing (*AI = .97*) and Journal of Marketing Research (*AI = .97*).

1. **Technical project requirements**

For our research project we analyze text data of 60 GB. We use advanced Natural Language Processing techniques including text analysis with dictionaries and seeded Latent Dirichlet Allocation (LDA). To conduct the analysis, we expect to require a substantial amount of RAM, of around 256 GB. Typically, we will use 32 CPU cores to employ parallelization, and we anticipate a computing time of around 48 hours per run. Our intermediate outputs are rather small and of approximately 5 GB in total. The computations take place in R and use a collection of different programming packages. We plan to only perform a few runs with the full dataset and are going to do the initial testing on a smaller sub-sample which would require fewer resources. We plan to finish our initial computations within the next 4 months. We do not foresee to require additional assistance in setting up the project.