CSR Communication on Twitter - A Scoping Review on Social Media Mining and Analytic Methods

Katharina Pilgrim
Witten/Herdecke University
katharina.pilgrim@uni-wh.de

Jonathan Koss
Witten/Herdecke University
jonathan.koss@uni-wh.de

Sabine Bohnet-Joschko
Witten/Herdecke University
sabine.bohnet-joschko@uni-wh.de

Abstract

Adopting corporate social responsibility (CSR) is becoming increasingly mandatory as international legislation puts pressure on companies to implement and report on appropriate CSR measures. As of 2024, a significant number of companies will need to report on CSR topics for the first time. To identify relevant topics that resonate best in the industry or even with one's own stakeholder groups and should therefore be picked up, addressed and reported on preferentially, social media mining (SMM) can be an efficient approach for companies. By reviewing applied SMM and analytic methods of Twitter data, we identified four methodological approaches that use algorithms to identify relevant CSR topics for companies to engage with. This scoping review thus provides a systematized overview of SMM pipelines for use, being equally relevant for academics and practitioners aiming at computational analysis of Twitter content regarding CSR activities and communication.

Keywords: Social Media Mining, CSR, Topic Model, Network Analysis, Sentiment Analysis

1. Introduction

1.1. Twitter data for CSR strategy success

Over 20.000 companies listed on an European-market will be required to implement and report on corporate social responsibility (CSR) activities for the first time from 2024 onwards due to the new Corporate Sustainability Reporting Directive (CSRD). Knowing which activities work best for stakeholders or within ones industry, and should therefore be picked up, addressed and reported on preferentially, becomes a key challenge. Utilizing existing publicly available information from sources with CSR topic engagement can provide a powerful solution (Wang & Huang, 2018). In this context, social media platforms offer large data on public CSR discussions. Twitter in particular provides a suitable environment for CSR communication (Zhang et al., 2011) as a medium for large user interaction on ethical corporate practices (Maiorescu-Murphy, 2022). To measure success and success factors of CSR communication on Twitter, researchers and practitioners try to employ all metrics the platform has to offer – by themselves or to create new ones. Stakeholder engagement for example can be measured on Twitter by the number of comments, likes and (textual) retweets per post (Giacomini et al., 2021a; Zhou et al., 2022). Adding the analysis of the user opinion expressed within a tweet, reply or retweet, by mining user sentiment (Pak & Paroubek, 2010) enables companies to understand factors facilitating positive user mentions (Alboqami et al., 2015; Chu & Sung, 2015) and to support topic selection suitable for the (online) community. While some find that retweets and likes are positive countersignals that reflect to what extent the message resonates with online stakeholders (Chae, 2021) others argue that they are indicators for increased customer loyalty (Iglesias et al., 2020), corporate reputation (O'Brien et al., 2018) or even the impact on society (Alboqami et al., 2015; Huete-Alcocer, 2017).

It is relevant for companies to understand the possibilities of collecting and analyzing Twitter data for strategy development and evaluation in the context of CSR measures (and reporting). As the sheer volume of such potentially valuable information makes its manual analysis highly inefficient and nearly impossible, artificial intelligence-based methods, such as social media mining (SMM) and analytics (the automated extraction and analysis of social media data) can be an effective alternative for companies to employ (Koss et al., 2021).

1.2. Research aim and question

As scoping reviews have become an increasingly popular approach for synthesizing research evidence by mapping the existing literature in the field of interest in terms of the volume, nature, and characteristics of the primary research (Pham et al., 2014), we aim to identify existing SMM and analytic approaches for research on CSR on Twitter employing a scoping (methods) review.

URI: https://hdl.handle.net/10125/102899
978-0-9981331-6-4

We chose Twitter as it offers low-threshold access to historical data via the publicly available API. Thus, the number of studies on CSR communication on Twitter, as a sentiment and opinion-forming platform, exceeds those that include other business networks, such as LinkedIn, by approximately three times (GoogleScholar hits of CSR and Twitter: 208,000 in August 2022 compared to 76.000 hits of CSR and LinkedIn). By identifying and explaining existing methodologies, and by grouping areas of application and specific software packages used, we promote their potential use by companies to address their own strategic CSR issues. This review contributes to the body of evidence on the possible applications and limitations of algorithm-based methods to scrape, mine and analyze Twitter data on CSR issues – and if reported on possible metrics to evaluate results. Thus, our research question is: *Which social media mining and analytic methods are used for what use cases (research aims) related to CSR on Twitter?*

2. Background on CSR reporting and standards

Companies across all industries are increasingly striving to consider environmental and social aspects within and beyond the scope of their business activities (Babiak & Trendafilova, 2011; Schaltegger & Burritt, 2017). The fact that companies not only pursue short-term profit maximization but also make their contribution to social development beyond products, jobs, and tax payments has a long tradition and has been referred to as corporate social responsibility (CSR) since the 1950s (Lantos, 2001). Companies began to understand CSR disclosure as an instrument to enhance the stakeholders' perception of its actions and ethics (Nekhili et al., 2017). Still, in 2011 the EU Commission found that only 1,000 of the 42,000 largest European companies disclose CSR related activities transparently and voluntarily. In response, Directive 2014/95/EU - also called the Non-Financial Reporting Directive (NFRD) - was adopted in 2014 (Fiandrino & Tonelli, 2021). For the first time, specific and mandatory requirements regarding the annual non-financial activity disclosure were established for certain large companies (>500 employees) in Europe. It introduced a 'double materiality perspective', meaning that companies have to report on how sustainability issues affect their business and on their own impact on people, and the environment. As a result, around 11,700 companies across Europe were required to prepare a sustainability report as of 2017 (EC, 2021a).

Besides the insufficient number of participating companies, one of the biggest problems regarding CSR reporting is mainly the report quality (Michelon et al., 2015). To address this issue, international standards have been developed. To understand the idea of these standards, we look at the beginnings of CSR on the international political platform. In 1992, the United Nations defined three pillars of sustainability at the Rio Conference, also referred to as sustainable triple-bottom-line (Purvis et al., 2019). It describes a balance of social, economic, and environmental goals that must be pursued simultaneously and reconciled to ensure sustainable development (Tomislav, 2018). For the first time, the importance of economic gains and societal benefits formally became the guiding principle of international policy. In 2015, United Nation (UN) member states committed to achieve 17 Sustainable Development Goals (SDGs), on a national level, which cover aspects of the sustainable triple-bottom line, as part of the subsequent Agenda 2030 (Caiado et al., 2018).

Thus, to support firms in a high quality and reliable sustainability reporting, developed standards try continuously to draw on international legal frameworks such as the SDGs to enable transparent, comparable and comprehensive reports (Tsalis et al., 2020), e.g. as an analytical measurement of negative and positive impacts within the value chain (Muñoz-Torres et al., 2018; Zimon et al., 2020). One relevant example for a CSR reporting standard provider is the Global Reporting Initiative (GRI). Founded in 1997, it serves as an internationally independent organization up to today whose self-proclaimed goal is to enable a comparable and standardized presentation of the environmental, social, and economic activities of large corporations, small and medium-sized enterprises (SMEs), other organizations and governments (de Villiers et al., 2022). Due to the early cooperation with the UN Global Compact, which recommends the application of the GRI standards to its members, it has established itself for the preparation of sustainability reports globally. GRI has also partnered with the European Financial Reporting Advisory Group (EFRAG), responsible for developing CSRD standards. GRI standards thus support the report preparation in accordance with the EU Directive, based on the UN SDGs.

Today, there is thus ample evidence that the information companies disclose, even by following standards, is still insufficient, resulting in yet another reinforcement of existing European regulations. In April 2021, the EU decided on a significant NRFD expansion, the so-called Corporate Sustainability Reporting Directive (CSRD), covering nearly all relevant reporting aspects: User group, independent audit requirements as well as reported content. For example, all companies listed on an EU-regulated market with a minimum of 250 or more employees will be subject to the new reporting requirements, affecting then at least 49,000 companies. Adding to the mandatory audit, sanctions will be imposed in the event of violations (EC, 2021b).

In addition to the associated resources investment following the increasing complexity of requirements, e.g., for risk management, mandatory sustainability reporting also offers opportunities. It enables companies to increase corporate reflectiveness and ensures sufficient transparency of the sustainability performance vis-à-vis key stakeholder groups (Wallage, 2000). As sustainability information will increasingly form an essential foundation for decision-making among investors (Grayson & Hodges, 2017), it is also associated with increases in firm valuations (Hamed et al., 2022; Ioannou & Serafeim, 2017). At the same time, even in markets without CSR disclosure regulations, there is an increased likelihood of voluntarily adopting reporting guidelines as firms seek the qualitative properties of comparability and credibility (Ioannou & Serafeim, 2017). Finally, high quality and reliable public reporting supports creating a culture of greater public accountability.

3. Research method

3.1. Data collection

We selected five databases with complementary research areas for a scoping review: (1) EconBiz, as the virtual subject library for the field of economics, (2) ScienceDirect, covering topics of physical sciences and engineering, life sciences, health sciences, social sciences and humanities, (3) Taylor and Francis Online, adding a source for information science, mathematics and statistics, (4) MDPI (Multidisciplinary Digital Publishing Institute), focusing on open access journal publication, and (5) PubMed for biomedical literature from MEDLINE. We chose to limit publications that date back until 2013 in view of the absolute monthly Twitter users as well as the monthly platform growth rate: in the first 7 years (until 2013), Twitter was able to surpass 200 million monthly active users worldwide, the growth rate has since then leveled off. At the same time, Twitter Inc. became a public trade company in November 2013, which added to a different platform use and is thus another reason to define 2013 for cut-off. All three authors as part of preliminary exploratory feasibility searches identified the keywords selected independently. Since machine learning approaches were rarely research topics, we decided not to include related keywords in the search strategy, but to define their application as a relevant inclusion criterion in the following. In an iterative process, keywords were extracted from relevant studies and condensed to define the two content areas (A) communication channel (*Twitter, tweeting*) and (B) communicated content (*CSR, Corporate Social Responsibility, Sustainability, Corporate Citizenship*). A correct database-specific application of meaningful search operators like Boolean or proximity operators was applied. The 187 publications as a result of applied search strategy, were systematically reviewed according to the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Page et al., 2021).

First, 45 duplicates were removed. We then drew up a list of criteria for inclusion and exclusion of articles in our review throughout the title (106), abstract and full-text screening (12). We included articles if all the following criteria were met: (i) the articles scope includes the target group relevant to this research (companies) or their direct stakeholders, (ii) the study employs machine learning approaches either stand-alone or in combination with other quantitative or qualitative methods, (iii) the article was published in a peer-reviewed publication and (iv) the article was available in English. We excluded articles according to the following criteria: (i) the article did not focus on Twitter but Social Media in general, (ii) the article focused on Twitter as a company and its CSR activities, (iii) the studies used additional data in analysis that is not extracted from Twitter, (iv) the studies used non-computational methods for analysis, (v) the studies relied on commercial preprocessing and analysis tools with no further reasonable explanation. Based on defined criteria, we excluded 106 publications in an initial title and subtitle screening. A subsequent full abstracts screening led to a selection of 27 publications for full text screening. Finally, we identified 15 relevant publications for analysis.

3.2. Data analysis

For each article, we collected data on (i) the research aim, (ii) the year in which the study was carried out, (iii) the geographical focus, (iv) the target group, (v) sample size, (vi) data sources – of which results are presented in chapter 4.1 as sample descriptives - methods for: (vii) SM scraping / data extraction and sample selection criteria (chapter 4.2), (viii) pre-processing, (ix) analysis, and (x) evaluation (chapters 4.3 & 4.4). To analyze the collected information, we used a narrative review synthesis to capture the broad range of research studying Twitter for CSR communication in our scoping review (Edo-Osagie et al., 2020). For structuring methodical approached, we employed pipelines for SMM by Koss et al. (Koss et al., 2021). The analysis process was chronologically as follows to ensure the requirements of transparency and reproducibility (Endenich & Trapp, 2020; Fatima & Elbanna, 2022): First, one of the authors read all of the articles and coded them based on the coding manual. We ensured inter- and intracoder reliability via multiple codings by the same author and by double coding of an article selection by a second author who was not part of the screening process. Deviations were subject to discussion until the authors agreed on which coding was appropriate and consistent with the remaining codings.

4. Results: SMM and analytic methods for CSR Communication on Twitter

4.1. Sample description

The mode publication year for our 15 articles was 2021 with a range of 2017 to 2022 after applying in- and exclusion criteria and an initial publication time limitation of 2013. Besides eight studies having no specific geographical focus, we found the representation of three countries in our sample: Italy (3) United States of America (USA) (3) and United Kingdom (1). We have 12 studies that are including company Twitter accounts, with one study especially focusing on special CSR accounts and three even disclosing the companies' names. 12 studies include tweets from company stakeholders or Twitter users in general (10), with one including CEO and another one NPO accounts. Thus, half of the sample includes more than one target group (8). Further, five studies focus on a specific industry: energy, mining, container shipping, news media and food. The final sample sizes regarding included Twitter accounts is not disclosed in 3 (of the included 15 studies) and ranges from 8 to 223.476 accounts in the other 12 studies (with a mean of 18.784 and median of 82). Sample size regarding Tweet count includes number of retweets and replies (when applicable) and ranges from 6.666 to 2 million Tweets (and a mean value of 355.311), with 3 studies not disclosing.

4.2. SM Data extraction and sample selection approaches

We found that research on CSR employing Twitter data and SMM and analytic methods focuses on two distinct areas: (A) CSR content exploration and (B) stakeholder reaction. Reviewing employed SM Scraping methods for data extraction and the associated definition of sample inclusion criteria in this context, we find them to be unspecific in both areas (A and B) and rather depending on the respective scope. In general, to scrape data, the Twitter API accessed via custom python scripts is used (B. K. Chae & E. O. Park, 2018; Mazza et al., 2022). Two studies employ specific software: Netlytic (Pilař et al., 2019) and Brandwatch (Jiang & Park, 2022). To extract data concerning the area of interest, e.g., a certain industry, data needs to be filtered. There are two approaches to define necessary sample inclusion criteria: (1) by the application of relevant CSR search terms (B. K. Chae & E. O. Park, 2018; Giacomini et al., 2021b; Jiang & Park, 2022; Pilař et al., 2019; Pons et al., 2021) e. g. #CSR, that are identified individually (Etter, 2014) or rely on or expand (Patuelli et al., 2021) existing keyword lists based on the GRI standards (Elkington, 1997; Gamerschlag et al., 2010; Jiang & Park, 2022) or (2) by applying a priori account selection criteria (Araujo & Kollat, 2018; Mazza et al., 2022; Patuelli et al., 2021; Salvatore et al., 2020; Xu & Saxton, 2019; Zhou et al., 2022). More specifically, we find five studies relying on selection according to company size (and geographical focus): Dow Jones Industrial Average Index (USA), S&P 1500 (USA), FTSE 350 (UK), Fortune 200 and 500. Two draw on the companies standing in the CSR domain to ensure CSR activity on Twitter in the first place: CR magazine' 100 best Corporate Citizen, AIDA (Italy). In addition, the Hootsuite for CEO (global) and Council on foundation List Website (USA) for NPO selection were employed.

4.3. SMM and Analytic Methods for content exploration (A)

Seven studies explore (A) CSR content exclusively. We were able to identify two generic pipelines regarding CSR content analysis, applying different unsupervised learning algorithms such as structural topic model (evolution of LDA considering meta data) (B. K. Chae & E. O. Park, 2018; Salvatore et al., 2020) or (semantic) network analysis (Patuelli et al., 2021; Pilař et al., 2019) (Fig. 1). Descriptive analyses such as hashtag frequencies often complement mentioned machine learning approaches (Mazza et al., 2022; Pilař et al., 2019).

4.3.1 Network Analysis Pipeline (NAP). Hashtags represent the main features in network analysis (Pilař et al., 2019). If extracted data is based on company account names, the data is then filtered by relevant CSR terms used as hashtags (Patuelli et al., 2021). Thus words without the symbol # are removed (Pilař et al., 2019). To enable meaningful analysis, data preprocessing is necessary. Data preprocessing includes consolidation of hashtags with the same meaning, but differing due to typos or punctuation (Patuelli et al., 2021). Inexact string matching can be used to aggregate those hashtags, by applying edit distance, e.g., by employing the *py_stringmatching python module* (Patuelli et al., 2021). Furthermore, preprocessing can involve splitting strings of hashtags with missing blank, e.g., "#CSR#substainabilty" (Pilař et al., 2019). To avoid duplicates, caused by capitalization, consistent lowercasing is applied (Pilař et al., 2019). Network features can then be generated in 3 ways: the networks nodes can be represented by (i) hashtags linked due to the co-occurrence frequencies (monopartite network) (Pilař et al., 2019), by (ii) (company) accounts, linked via co-occurring hashtag usage, indicating similar content (bipartite network) (Patuelli et al., 2021), or by (iii) displaying both, connections between companies or companies and hashtags

(Giacomini et al., 2021b). All approaches can be displayed as undirected graphs (Giacomini et al., 2021b; Patuelli et al., 2021; Pilař et al., 2019). The (i) monopartite network can be used to analyze which hashtags are often communicated together, while the (ii) bipartite network allows the identification of companies communicating similar content, represented by hashtags. For implementation of the Bipartite configuration Model (BiCM) for example, the *python module NEMtropy* can be used (Patuelli et al., 2021) (Vallarano et al., 2021). In sum, both approaches offer the possibility to derive CSR topics by community identification, e.g., by applying the *Louvain algorithm* (Patuelli et al., 2021; Pilař et al., 2019).

To label CSR topics most important hashtags in a cluster, which can be identified by analyzing frequency, eigenvector centrality or degree centrality, need to be analyzed (Patuelli et al., 2021; Pilař et al., 2019). To avoid the identification of local communities and thus increase the focus and quality of community detection, the removal of infrequent nodes (hashtags) can be considered (Pilař et al., 2019). Finally, to evaluate the quality of community detection the modularity metric Q can be applied (Pilař et al., 2019).

al., 2020). Additional processing the textual data as a lexical feature through an approach to lemmatization called Part of Speech can be applied, e.g., by annotation of language elements in the texts with a languageagnostic tool, the *R package UDPipe* (Straka & Straková, 2017) (Mazza et al., 2022). Which preprocessing steps are necessary, needs to be evaluated for each approach individually. When applying topic models in short text such as Twitter posts, specific additional preprocessing steps can be considered such as grouping tweets by user to reduce data sparsity or removing tweets which include less than a certain number of words (B. K. Chae & E. O. Park, 2018). To remove non-English Tweets, a language identification Python package such as *langid* can be added (B. Chae & E. Park, 2018).

For analysis either LDA topic model or Structural Topic Model (STM) employing the *stm: R Packag* (B. Chae & E. Park, 2018; Salvatore et al., 2020) (Roberts et al., 2019) can be applied.

For evaluation and to obtain optimal topic number, researchers apply different metrics such as topic coherence and/or topic exclusivity (B. K. Chae & E. O. Park, 2018; Salvatore et al., 2020) or density based measures. The topics content always needs manual labeling by analyzing the most important words in each topic (B. K. Chae & E. O. Park, 2018; Salvatore et al., 2020). Results can be described by topic labels, topic prevalence (Mazza et al., 2022) over time (B. K. Chae & E. O. Park, 2018; Salvatore et al., 2020)) and correlations, often presented as topic networks (B. K. Chae & E. O. Park, 2018; Mazza et al., 2022; Salvatore et al., 2020).

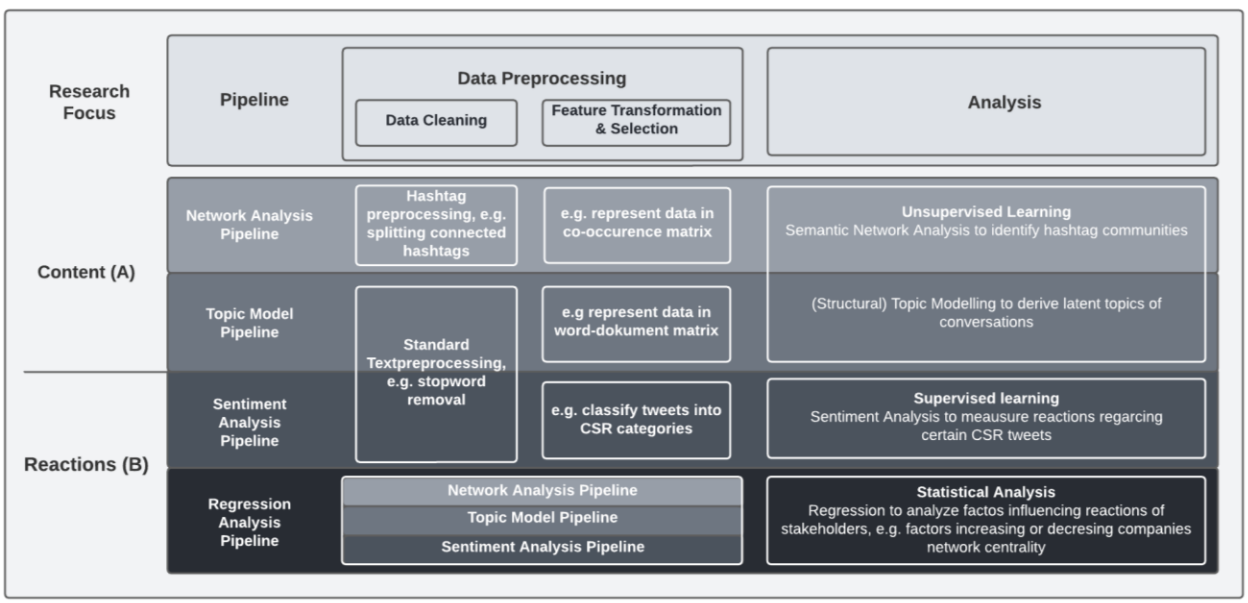

Figure 1. SMM and analytics pipelines for CSR research on Twitter.

4.3.2 Topic Model Pipeline (TMP). In contrast to network analysis using hashtags as features, words represent the main feature in topic modelling. Text preprocessing involves removing hyperlinks, usernames or numbers from the dataset (B. K. Chae & E. O. Park, 2018). Furthermore elimination of stop words, punctuation or infrequent words as well as stemming and lowercasing can be applied (Mazza et al., 2022; Salvatore et

4.4. SMM and Analytic Methods for Stakeholder Reaction (B)

Eight studies explore (B) the reactions of Twitter users or companies to CSR content (Fig. 1) (Giacomini et al., 2021b; Grover et al., 2019; Pons et al., 2021; Zhou et al., 2022) or communication strategies (Araujo & Kollat, 2018; Jiang & Park, 2022; Xu & Saxton, 2019) (Gregory D Saxton et al., 2021). These approaches also include the caption of content for example on which SDGs is reported (Grover et al., 2019; Zhou et al., 2022) or determination of communication strategy in a first step (Gregory D Saxton et al., 2021). To identify whether a tweet contains CSR information, a variety of methods is applied, similar to what we described for the TMP. One study for example employed the *Twitter-LDA algorithm* by Zhoa et al. (Zhao et al., 2011) to categorize topics within documents (Crowley et al., 2019). Twitter-LDA extends the basic LDA model to work with shorter documents of at most 140 characters and this adapted to Twitter data characteristics, by incorporating correlations between words across Twitter users (Crowley et al., 2019). To analyze user reaction to content, which can also be referred to as opinion mining (Pak & Paroubek, 2010), researchers either capture the keyword sentiment (Pons et al., 2021) or sentiment of CSR related tweets (Giacomini et al., 2021b). Two studies don't fit in either of the identified pipelines for content reaction: one study thus measured content reactions explicitly by behavioral features such as number of likes and retweets (Patuelli et al., 2021). The other one employed grey relational analysis with identifying stakeholder mean sentiment employing *VADER sentiment analysis* (among other variables), which relies on a dictionary that maps lexical features to emotion intensities known as sentiment scores (Elbagir & Yang, 2019; Zhou et al., 2022), to then use it as one of four variables for the decision matrix (Zhou et al., 2022). To analyze reactions to communication strategies (Araujo & Kollat, 2018; Jiang & Park, 2022; Xu & Saxton, 2019) or to determine how companies react to user content (Gregory D Saxton et al., 2021) researchers apply regression analysis (Araujo & Kollat, 2018; Jiang & Park, 2022; Gregory D Saxton et al., 2021; Xu & Saxton, 2019). Depending on the aim of analysis, pipelines differ fundamentally.

4.4.1 Sentiment Analysis Pipeline (SAP). Three studies measured reactions to CSR content by analyzing sentiment of keywords (Pons et al., 2021), companies of a certain sector (Zhou et al., 2022) or CSR related tweets (Giacomini et al., 2021b). First, data needs to be preprocessed, e.g., by removing stop words, URLs and whitespaces (Giacomini et al., 2021b; Pons et al., 2021). Then, the content of which reactions in means of sentiment should be measured, needs to be identified (Giacomini et al., 2021b; Pons et al., 2021). For instance, Giacomini et al. classify tweets into CSR categories by filtering for CSR related keywords (Giacomini et al., 2021b). Next, the sentiment of certain CSR content can be determined to analyze public reaction to those posts, or certain CSR actions (Giacomini et al., 2021b; Pons et al., 2021). In general, sentiment analysis can be carried out by using pretrained (Giacomini et al., 2021b) or customized models (Pons et al., 2021), leading to additional preprocessing steps, including manual curation of the training data sets (Pons et al., 2021). One example of a pretrained model is the application of the *algorithm SentITA*, exclusively devoted to the Italian language and based on long short-term memory (LSTM) networks with an attention mechanism (Giacomini et al., 2021b). Pons et al. train a *Naive Bayes classifier* for sentiment classification (Pons et al., 2021). To capture reactions, sentiment analysis is often complemented by descriptive statistics of using features such as likes or quotes of CSR related tweets (Giacomini et al., 2021b).

In general, to evaluate any analytical steps involving manual coding, intercoder reliability, e.g., using Krippendorff's α (Jiang & Park, 2022), is employed (Araujo & Kollat, 2018). As classification performance can directly depend on the quality of features obtained from the training dataset, for evaluation of balanced data, the trained model accuracy needs to be tested (Pons et al., 2021). For unbalanced data the F-score is used to evaluate results (Araujo & Kollat, 2018). Only one study exclusively belonging to the SAP, refers to results of another study regarding the Naive Bayes algorithm accuracy, but does not report any scores (Pons et al., 2021).

4.4.2 Regression Analysis Pipeline (RAP). Five studies (Araujo & Kollat, 2018; Grover et al., 2019; Jiang & Park, 2022; Xu & Saxton, 2019) (Gregory D Saxton et al., 2021) adopt a mixed research methodology, using selective methods from SMM and analytics and combining them with conventional data analysis approaches from social science research methodologies (Grover et al., 2019). They examine the influence of communicated topics (Jiang & Park, 2022; Xu & Saxton, 2019) or communication strategies on public responses using multiple regression (Araujo & Kollat, 2018), set of ordinary least squares (OLS) linear regression models (Xu & Saxton, 2019), logistic regression (Gregory D Saxton et al., 2021) or path analysis (Jiang & Park, 2022). Pipelines differ in considered features, depending on the research question and the applied regression method. Variable generation for the regression model chosen can include one or a combination of the described Network Analysis Pipeline (NAP), Topic Model Pipeline (TMP), or Sentiment Analysis Pipeline (SAP). For example, to identify CSR topics, tweets can be assigned to a predefined CSR category system with the help of a created codebook (e.g., based on the UN SDGs), as a dictionary method and thus specific form of content/text analysis (Grover et al., 2019). Another approach employed the multinomial Naive Bayes algorithm to automatically categorize all tweets included in the sample relative to their CSR content (Araujo & Kollat, 2018). For evaluation accuracy and F-score is reported (Araujo & Kollat, 2018).

Further, companies communication behavior was captured by including behavioral information as independent variables, for instance number of replies, mentions (Araujo & Kollat, 2018; Jiang & Park, 2022), retweets (Jiang & Park, 2022), and tweets (Xu & Saxton, 2019). Storytelling is another example for an independent variable. To detect the elements of storytelling within a tweet, one study used Linguistic Inquiry and Word Count (LIWC) software to categorize tweets and measure the tweets emotions (by the presence of affective words) or aspirational talk (by the presence of words relating to future focus) (Araujo & Kollat, 2018). Studies also measured reactions to strategies by defining different dependent variables e. g. by (i) applying sentiment analysis using Brandwatch software (Jiang & Park, 2022), or (ii) with a supervised machine learning technique (hand-coding 1,000 tweets to train an SVM model) (Gregory D Saxton et al., 2021), by (iii) capturing stakeholder support considering the total number of retweets (Araujo & Kollat, 2018; Jiang et al., 2020), or (iv) likes (Araujo & Kollat, 2018). Studies, that applied network analysis (Jiang & Park, 2022; Xu & Saxton, 2019) for the generation of dependent and independent variables, determine centrality for example as a measure of a company's popularity in the Twitter network (Jiang & Park, 2022) or in-degree centrality of the stakeholder message sender in a chosen network as a measure of stakeholder power using the social networking analysis Python package *Networkx* (Gregory D. Saxton et al., 2021). To control for confounding factors, known control variables are included.

The quality of the regression model can be indicated by the adjusted r square.

5. Discussion

Results show that SMM and analytic methods have gained prominence specifically since the 2020s for doing research on Twitter to answer questions on CSR. While previously it were mostly manual methods for data extraction, preprocessing and analysis, we see that increasingly automated computerized methods are used along the whole pipeline. In this way, limitations such as small sample sizes as well as merely a priori defined categories that do not provide new insights are worked around. Moreover, depending on the primary goal, the exploration of CSR topics exclusively or the analysis of reactions to CSR topic, communication strategy or even tweet structure prioritization, we were able to differentiate the identified SM analytics approaches into four pipelines - two for each goal. Topic models or network analyses are used for the former, each of which employs data preprocessing methods that are rarely presented in detail - cross-referencing to other studies is often used here. Sentiment analyses as well as a mixed-methods approach employing a combination of SMM and analytics (topic model, network and sentiment analyses) with classical statistical methods such as regression or path analysis are employed for analyzing user reactions. Here, the outputs generated by means of SM analytics predominantly describe dependent variables, but in some cases also independent ones to test hypotheses. On a methodological level, we find no framework that allows studies using SM analytic methods alone, in combination with each other, or with conventional data analysis approaches from social science to be situated in quantitative, qualitative, or mixed-method as is the case with classical study designs. Future research can start at this point and develop the basis, frameworks and systematization to build up and classify the so far very heterogeneously used labels and descriptions of study designs with SMM and analytics processes in a comparable way.

Although we have succeeded in identifying generic pipelines for different questions, our study is subject to the usual limitations of scoping reviews (Tricco et al., 2016). We can thus not preclude the possibility of other (automated) SMM and analytic methods being used for the analysis of Twitter data on CSR topics. Looking at methodological pipeline overviews in other fields, such as political communication, Stieglitz and Dang-Xuan (Stieglitz & Dang-Xuan, 2013) for example further identified trend analysis as one specific form of topic analysis. However, in the context of the presented systematic literature review process for the use case CSR, the pipelines present all identified methods. Further, we find limitations in each study, which thus also add further limitations to our review results in terms of generalizability and comparability. Nine general limitations are discussed in the following: (i) identified pipelines vary greatly in the usage and designation/assignment of individual parameters such as count of likes or retweets - as there is no universally applicable systematization or assignment as to which variables are best for measuring, e.g., qualitative constructs such as corporate reputation. (ii) Only some studies, which respectively rarely have a method focus, name employed software or disclose codebooks for training the algorithms and present use case specific pipelines. (iii) The regressions performed sometimes only explain little of the variances, e.g. when investigating the impact of board diversity on CSR communications on Twitter, which raises the question of whether Twitter is the right source to answer these questions. Overall, when making statements on the significance of CSR activities on Twitter for decision-making, it should always be noted that Twitter's users do not represent the general global population, being younger, more left-leaning and more affluent. In addition, social bots can bias results. As there was no evidence on approaches to deal with this problem, it should be considered in future deployments.

Surprisingly, (iv) hardly any study even addresses limitations of the employed method. In addition, (v) the majority of studies do not provide any information on the algorithm quality, i.e., they miss to include relevant key metrics to evaluate algorithms performance. In some cases, when reported, they can even be misleading, for example, applying accuracy on unbalanced data. Furthermore, it is (vi) rarely clear how many rounds of manual coding with how many different datasets of what sample size were performed. To evaluate algorithm quality and thus the reliability of the results, it is important to pay special attention to this in future studies. Further, (vii) dictionaries of CSR keywords used for sample extraction or to identify relevant CSR content as part of the analysis are limited and inconsistent. Open access solution of curated CSR keyword sets for the comparability of studies with, e.g., national focus or also to capture all relevant content would be desirable, building on SDG or GRI keywords. One general limitation (viii), especially when it comes to user reactions and sentiment analysis, is the use of algorithms to identify irony or sarcasm. Only one study in our sample uses SVM as the most commonly used algorithm found in the literature for detecting sarcasm on Twitter, others do not address this or use the workaround of hashtag analysis. Similarly, (ix) emoticons, gifs, or memes, which are often part of speech on Twitter, are not included in the analysis. Additional image classifiers used in the pipelines could provide added value here.

6. Conclusion

Given the growing involvement of international companies in CSR activities, driven not least by increasing political pressure and the introduction of mandatory sustainability reporting, Twitter is proving to be a relevant resource for analysis in the course of a company's CSR (communication) strategy development and at the same time as a channel for implementing this strategy. Our review results contribute to the understanding of how Twitter data can be accessed and employed for strategy development and evaluation in the context of CSR measures by identifying four generic approaches to shed light on two distinct questions: (1) what CSR topics are communicated on Twitter (within my industry, by my competitors or my community) and (2) what topics or communication approaches resonate best within my desired network or selected stakeholders? Results provide insides to four different SMM and analytic approaches: (i) topic models and (ii) network analysis for topic exploration as well as (iii) sentiment analysis and (iv) regression analysis (in which variables can be developed using SMM and analytic approaches) to analyze and evaluate stakeholder reactions. The identified pipelines can facilitate their use and support managers in implementing SMM for CSR strategy development and decision making.

7. References

Alboqami, H., Al-Karaghouli, W., Baeshen, Y., Erkan, I., Evans, C., & Ghoneim, A. (2015). Electronic word of mouth in social media: the common characteristics of retweeted and favourited marketer-generated content posted on Twitter. *International Journal of Internet Marketing and Advertising, 9*(4), 338-358.

Araujo, T., & Kollat, J. (2018). Communicating effectively about CSR on Twitter: The power of engaging strategies and storytelling elements. *Internet Research*.

Babiak, K., & Trendafilova, S. (2011). CSR and environmental responsibility: Motives and pressures to adopt green management practices. *Corporate social responsibility and environmental management, 18*(1), 11-24.

Caiado, R. G. G., Leal Filho, W., Quelhas, O. L. G., de Mattos Nascimento, D. L., & Ávila, L. V. (2018). A literature-based review on potentials and constraints in the implementation of the sustainable development goals. *Journal of cleaner production, 198*, 1276-1288.

Chae, B., & Park, E. (2018). Corporate Social Responsibility (CSR): A Survey of Topics and Trends Using Twitter Data and Topic Modeling. *Sustainability, 10*(7), 2231. https://www.mdpi.com/2071-1050/10/7/2231

Chae, B. K., & Park, E. O. (2018). Corporate social responsibility (CSR): A survey of topics and trends using Twitter data and topic modeling. *Sustainability, 10*(7), 2231.

Chae, M.-J. (2021). Driving consumer engagement through diverse calls to action in corporate social responsibility messages on social media. *Sustainability, 13*(7), 3812.

Crowley, R. M., Huang, W., Lu, H., & Luo, W. (2019). Do firms manage their CSR reputation? Evidence from Twitter. *Evidence from Twitter (October 20, 2019)*.

de Villiers, C., La Torre, M., & Molinari, M. (2022). The Global Reporting Initiative's (GRI) Past, Present and Future: Critical reflections and a research agenda on sustainability reporting (standard-setting). *Pacific Accounting Review*(ahead-of-print).

European Commission (2021a). *Corporate sustainability reporting*. Retrieved 13.02.2022 from https://ec.europa.eu/info/business-economy-euro/company-reporting-and

-auditing/company-reporting/corporate-sustainability-reporting_en
European Commission (2021b). *EC publishes proposed Corporate Sustainability Reporting Directive* https://www.iasplus.com/en/news/2021/04/csrd
Edo-Osagie, O., De La Iglesia, B., Lake, I., & Edeghere, O. (2020). A scoping review of the use of Twitter for public health research. *Computers in biology and medicine, 122*, 103770.
Elbagir, S., & Yang, J. (2019). Twitter sentiment analysis using natural language toolkit and VADER sentiment. Proceedings of the international multiconference of engineers and computer scientists,
Endenich, C., & Trapp, R. (2020). Ethical implications of management accounting and control: A systematic review of the contributions from the Journal of Business Ethics. *Journal of business ethics, 163*(2), 309-328.
Etter, M. (2014). Broadcasting, reacting, engaging - three strategies for CSR communication in Twitter [Aufsatz in Zeitschrift, Article in journal]. *Journal of communication management : an international journal*, *18*(4), 322-342.
Fatima, T., & Elbanna, S. (2022). Corporate Social Responsibility (CSR) Implementation: A Review and a Research Agenda Towards an Integrative Framework. *Journal of business ethics*, 1-17.
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*, *13*(2), 763.
Gamerschlag, R., Moeller, K., & Verbeeten, F. (2010). Determinants of voluntary CSR disclosure: Empirical evidence from Germany. *Review of Managerial Science, 5*.
Giacomini, D., Martini, M., Sancino, A., Zola, P., & Cavenago, D. (2021a). Corporate social responsibility actions and organizational legitimacy at the peak of COVID-19 : a sentiment analysis. *Corporate Governance: The International Journal of Business in Society*, *21*(6), 1043-1058. https://doi.org/10.1108/cg-09-2020-0389
Giacomini, D., Martini, M., Sancino, A., Zola, P., & Cavenago, D. (2021b). Corporate social responsibility actions and organizational legitimacy at the peak of COVID-19: a sentiment analysis. *Corporate Governance: The International Journal of Business in Society*.
Grayson, D., & Hodges, A. (2017). *Corporate social opportunity!: Seven steps to make corporate social responsibility work for your business*. Routledge.
Grover, P., Kar, A. K., & Ilavarasan, P. V. (2019). Impact of corporate social responsibility on reputation—Insights from tweets on sustainable development goals by CEOs. *International journal of information management, 48*, 39-52.
Hamed, R. S., Al-Shattarat, B. K., Al-Shattarat, W. K., & Hussainey, K. (2022). The impact of introducing new regulations on the quality of CSR reporting: Evidence from the UK. *Journal of International Accounting, Auditing and Taxation, 46*, 100444.
Huete-Alcocer, N. (2017). A literature review of word of mouth and electronic word of mouth: Implications for consumer behavior. *Frontiers in psychology, 8*, 1256.
Iglesias, O., Markovic, S., Bagherzadeh, M., & Singh, J. J. (2020). Co-creation: A key link between corporate social responsibility, customer trust, and customer loyalty. *Journal of business ethics, 163*(1), 151-166.
Ioannou, I., & Serafeim, G. (2017). The consequences of mandatory corporate sustainability reporting. *Harvard Business School research working paper*(11-100).
Jiang, H., Sun, L., Ran, J., Bai, J., & Yang, X. (2020). Community Detection Based on Individual Topics and Network Topology in Social Networks. *IEEE Access, 8*, 124414-124423. https://doi.org/10.1109/access.2020.3005935
Jiang, Y. N., & Park, H. (2022). Mapping networks in corporate social responsibility communication on social media: A new approach to exploring the influence of communication tactics on public responses. *Public Relations Review, 48*(1), 102143.
Koss, J., Rheinlaender, A., Truebel, H., & Bohnet-Joschko, S. (2021). Social media mining in drug development—Fundamentals and use cases. *Drug discovery today, 26*(12), 2871-2880.
Lantos, G. P. (2001). The boundaries of strategic corporate social responsibility. *Journal of consumer marketing*.
Maiorescu-Murphy, R. D. (2022). Business-centered versus socially responsible corporate diversity communication. An assessment of stakeholder (dis) agreement on Twitter. *Public Relations Review, 48*(1), 102138.
Mazza, R., Zavarrone, E., Olivieri, M., & Corsaro, D. (2022). A text mining approach for CSR communication: an explorative analysis of energy firms on Twitter in the post-pandemic era. *Italian Journal of Marketing*, 1-24.
Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine, 151*(4), 264-269.
Muñoz-Torres, M. J., Fernández-Izquierdo, M. Á., Rivera-Lirio, J. M., Ferrero-Ferrero, I., Escrig-Olmedo, E., Gisbert-Navarro, J. V., & Marullo, M. C. (2018). An assessment tool to integrate sustainability principles into the global supply chain. *Sustainability, 10*(2), 535.
Nekhili, M., Nagati, H., Chtioui, T., & Rebolledo, C. (2017). Corporate social responsibility disclosure and market value: Family versus nonfamily firms. *Journal of Business Research, 77*, 41-52.
O'Brien, I. M., Jarvis, W., Soutar, G., & Ouschan, R. (2018). Co-creating a CSR strategy with customers to deliver greater value. In *Disciplining the Undisciplined?* (pp. 89-107). Springer.
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*, 105906.
Pak, A., & Paroubek, P. (2010). Twitter as a corpus for sentiment analysis and opinion mining. Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10),
Patuelli, A., Caldarelli, G., Lattanzi, N., & Saracco, F. (2021). Firms' challenges and social responsibilities during

Covid-19: A Twitter analysis. *PLoS One, 16*(7), e0254748.
Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: advancing the approach and enhancing the consistency. *Research synthesis methods, 5*(4), 371-385.
Pilař, L., Kvasničková Stanislavská, L., Pitrova, J., Krejčí, I., Ticha, I., & Chalupova, M. (2019). Twitter analysis of global communication in the field of sustainability. *Sustainability, 11*(24), 6958.
Pons, A., Vintrò, C., Rius, J., & Vilaplana, J. (2021). Impact of Corporate Social Responsibility in mining industries. *Resources Policy, 72*, 102117.
Purvis, B., Mao, Y., & Robinson, D. (2019). Three pillars of sustainability: in search of conceptual origins. *Sustainability science, 14*(3), 681-695.
Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software, 91*, 1-40.
Salvatore, C., Bianchi, A., & Biffignandi, S. (2020). Communicating Corporate Social Responsibility through Twitter: a topic model analysis on selected companies. CARMA 2020: 3rd International Conference on Advanced Research Methods and Analytics,
Saxton, G. D., Ren, C., & Guo, C. (2021). Responding to diffused stakeholders on social media : connective power and firm reactions to CSR-related Twitter messages [Aufsatz in Zeitschrift, Article in journal]. *Journal of business ethics, 172*(2), 229-252. https://link.springer.com/content/pdf/10.1007/s10551-020-04472-x.pdf;https://doi.org/10.1007/s10551-020-04472-x
Saxton, G. D., Ren, C., & Guo, C. (2021). Responding to diffused stakeholders on social media: Connective power and firm reactions to CSR-related Twitter messages. *Journal of Business Ethics, 172*(2), 229-252.
Schaltegger, S., & Burritt, R. (2017). *Contemporary environmental accounting: issues, concepts and practice*. Routledge.
Stieglitz, S., & Dang-Xuan, L. (2013). Social media and political communication: a social media analytics framework. *Social network analysis and mining, 3*(4), 1277-1291.
Straka, M., & Straková, J. (2017). Tokenizing, pos tagging, lemmatizing and parsing ud 2.0 with udpipe. Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies,
Tomislav, K. (2018). The concept of sustainable development: From its beginning to the contemporary issues. *Zagreb International Review of Economics & Business, 21*(1), 67-94.
Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K., Colquhoun, H., Kastner, M., Levac, D., Ng, C., Sharpe, J. P., & Wilson, K. (2016). A scoping review on the conduct and reporting of scoping reviews. *BMC medical research methodology, 16*(1), 1-10.
Tsalis, T. A., Malamateniou, K. E., Koulouriotis, D., & Nikolaou, I. E. (2020). New challenges for corporate sustainability reporting: United Nations' 2030 Agenda for sustainable development and the sustainable development goals. *Corporate social responsibility and environmental management, 27*(4), 1617-1629.
Vallarano, N., Bruno, M., Marchese, E., Trapani, G., Saracco, F., Cimini, G., Zanon, M., & Squartini, T. (2021). Fast and scalable likelihood maximization for Exponential Random Graph Models with local constraints. *Scientific Reports, 11*(1), 1-33.
Wallage, P. (2000). Assurance on sustainability reporting: An auditor's view. *Auditing: A Journal of Practice & Theory, 19*(s-1), 53-65.
Wang, R., & Huang, Y. (2018). Communicating corporate social responsibility (CSR) on social media: How do message source and types of CSR messages influence stakeholders' perceptions? *Corporate communications: An international journal*.
Xu, W., & Saxton, G. D. (2019). Does stakeholder engagement pay off on social media? A social capital perspective. *Nonprofit and Voluntary Sector Quarterly, 48*(1), 28-49.
Zhang, M., Jansen, B. J., & Chowdhury, A. (2011). Business engagement on Twitter: a path analysis. *Electronic Markets, 21*(3), 161-175.
Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H., & Li, X. (2011). Comparing twitter and traditional media using topic models. European conference on information retrieval,
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*, 104853.
Zimon, D., Tyan, J., & Sroufe, R. (2020). Drivers of sustainable supply chain management: Practices to alignment with un sustainable development goals. *International Journal for Quality Research, 14*(1).