Cheap Talk: Topic Analysis of CSR Themes on Corporate Twitter

Nile Phillips[1], Sathvika Anand[1], Michelle Lum[1], Manisha Goel[2], Michelle Zemel[2], and Alexandra Schofield[1]
[1]Harvey Mudd College, Claremont, CA, USA
{nphillips, sanand, milum}@hmc.edu, xanda@cs.hmc.edu
[2]Pomona College, Claremont, CA, USA
{manisha.goel, michelle.zemel}@pomona.edu

Abstract

Numerous firms advertise action around corporate social responsibility (CSR) on social media. Using a Twitter corpus from S&P 500 companies and topic modeling, we investigate how companies talk about their social and sustainability efforts and whether CSR-related speech predicts Environmental, Social, and Governance (ESG) risk scores. As part of our work in progress, we present early findings suggesting a possible distinction in language between authentic discussion of positive practices and corporate posturing.

Keywords: corporate social responsibility, ESG, topic models, Twitter, social media, sustainability

1. Introduction

The last two decades have witnessed an urgent recognition by investors, and in response, firms, of a role for corporations in *corporate social responsibility* (CSR) (Bowen, 2013) and environmental stewardship. CSR integrates societal goals into firms' objectives, potentially channeling private investment towards a public good such as combating climate change and addressing inequality. Investors' demand for CSR activities has increased dramatically over this time period: the total market value of US assets managed with ESG strategies in 2020 totaled $17.1 trillion, a 33% increase from 2018's value and ake 25-fold increase relative to 1995 (US SIF Foundation, 2020). While company approaches to CSR may not impact their bottom line (McWilliams and Siegel, 2000), there are political, ethical, and social implications for why a company may build a CSR-focused strategy, with or without a profit motive (Garriga and Melé, 2004).

Social media platforms like Twitter[1] allow companies to communicate publicly about CSR to improve brand awareness and perception (Pilgrim and Bohnet-Joschko, December 2022; Araujo and Kollat, 2018). The embrace of Twitter as a platform to reach shareholders has included the creation of distinct corporate accounts, such as @KelloggsCompany or @CocaColaCo, focused not on products but corporate actions. Our ongoing project explores whether corporate Twitter messaging describes genuine commitments to social goals or instead is an example of "cheap talk" to paint companies in a positive light. We examine the behavior of S&P 500 English-language Twitter accounts, using a topic model to characterize themes in how they communicate about CSR. We present our work in progress, in which we find both concrete, action-oriented CSR-focused topics and more abstract topics highlighting sustainability and social good. We also compare our behavioral findings with Sustainalytics ESG risk scores to demonstrate that less concrete topics can correlate with increased risk.

[1]Our dataset predates renaming Twitter to "X" in 2023.

2. Background

2.1. Corporate Social Responsibility

Investor demand for firm CSR commitments can be explained by two dominant competing theories. Under the "doing well by doing good" theory, investor demand for integration of CSR stems from a belief that CSR activities lead to increased financial benefit to shareholders (McWilliams and Siegel, 2000; Orlitzky et al., 2003). In this theory, shifting to cleaner technologies, employing a diverse workforce, or partnering with local communities, for example, are long-term profit maximizing decisions. In contrast, an alternate theory suggests that demand for CSR is driven by non-pecuniary benefits to investors, such as cleaner air and social equality (Garriga and Melé, 2004). Under either theory, CSR creates value for investors and may drive engagement with current or prospective investors.

Through channels such as financial reports, shareholder calls, and more recently, social media, firms can signal their commitments to CSR to current shareholders, potential investors, consumers, and employees (Araujo and Kollat, 2018). To the extent that signaling a CSR commitment is less costly than executing on the commitment, especially in the less-regulated landscape of social media, the conditions for "cheap talk", or in the case of environmental initiatives, "greenwashing", exist. There is growing evidence that this phenomenon of "cheap talk" is present in social media discus-

Proceedings of the Joint Workshop of the 7th FinNLP, the 5th KDF, and the 4th ECONLP, pages 203–211
20 May, 2024.

sion of ESG commitments. Crowley et al. (2019) show that firms strategically use Twitter communications to "greenwash," i.e., exaggerate their CSR activities. In fact, those that are rated worse on ESG rankings talk more about their initiatives to build a more positive reputation even if this talk is only cheap and not consistent with their actions in reality. Baker et al. (2023) demonstrate that firms similarly use voluntary disclosures to make strong statements about their commitment to diversity initiatives but significantly lag in their actions. This helps build their reputations with customers and investors, and also improve their ESG ratings. Attig and Boshanna (2023) show, however, that such cheap talk worsens firms' market performance.

2.2. Twitter Analysis for CSR

Twitter data has been used for a variety of corporate analyses, including predicting stock behavior (Si et al., 2013, 2014) and financial stance detection (Conforti et al., 2022). Recent existing work also suggests that CSR communication is present on Twitter, including work from Pilgrim and Bohnet-Joschko (December 2022) surveying existing reported-on CSR strategies in digital media and Johnson and Greenwell (2022) analyze 200+ UK companies and the practice of greenwashing (defined as when a company presents itself as environmentally-friendly, even when its actions actually say otherwise), yielding no evidence for greenwashing across UK companies, but signs that environmental messaging occurs with low frequency on company Twitter accounts.

Rybalko and Seltzer (2010) and Okazaki et al. (2020) took a dialogic approach to analyzing CSR communications on Twitter by focusing on dialogue between brands and Twitter users, as encouraged in public relations literature (Kent and Taylor, 1998). Rybalko and Seltzer (2010) found that *Fortune 500* companies tend to underuse dialogue to engage their stakeholders, while Okazaki et al. (2020) found that companies mostly were not explicitly using CSR to engage on Twitter. These two works inform our strategy for examining our own corpus: we focus on company tweets that are not replies or retweets, and we use a many-topic topic model to try to access more diffused CSR-related themes.

Salvatore et al. (2022) use a structural topic model (a weakly supervised approach) to explore how businesses used social media to communicate CSR, specifically in relation to the Sustainable Development Goals set by the United Nations' 2030 agenda, using tweets from the 30 largest firms according to the Dow Jones Industrial Average in August 2020. While our findings echo the focus on social and environmental issues for these companies, we broaden our focus to S&P 500 companies and use an unsupervised topic model.

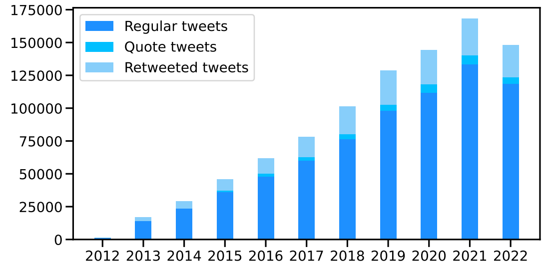

Figure 1: The composition of the initial dataset, broken down by tweet type and year, showing how API limits of 3,200 tweets per company reduced data availability for earlier years of our dataset.

3. Data

3.1. Collection Process

To gather Twitter handles for our companies, we scraped the websites of S&P 500 companies as listed on Wikipedia for all front-page links to Twitter handles. We augmented these Twitter handles with those listed in Twitter profiles for these companies. After manual vetting, we added obvious missing firm accounts, e.g. Match Group's subsidiaries. We excluded customer support Twitter accounts as well as regional accounts that were not immediately listed by companies on their website. With our list of S&P 500 Twitter handles, we used the Twitter API to retrieve as many tweets as possible from each company's Twitter account, going back at most 10 years from November 2022. Only data from more recent years was available for more prolific accounts due to the 3,200-tweet API limit for account history. Tweets in languages other than English were filtered out using the `fasttext-langdetect` library (Joulin et al., 2016b,a).

3.2. Composition

The initial dataset included 1,009,703 tweets from 536 distinct Twitter accounts. The dates of the tweets range from December 2012 to November 2022, with the breakdown of tweets by year shown in Figure 1. The parent companies of the Twitter accounts represent 11 distinct GICS Sectors, including Financials, Information Technology, Energy, Industrials, Consumer Staples, Health Care, Utilities, and Real Estate.

Tweets were tokenized using the Tweet Tokenizer from the Natural Language Toolkit (Bird et al., 2009), and tags to other users were replaced with "@TAG@" to prevent the formation of topics purely centered around tags. All terms were lowercased and stripped of trailing whitespace. Terms from the NLTK built-in English stoplist were filtered out in training.

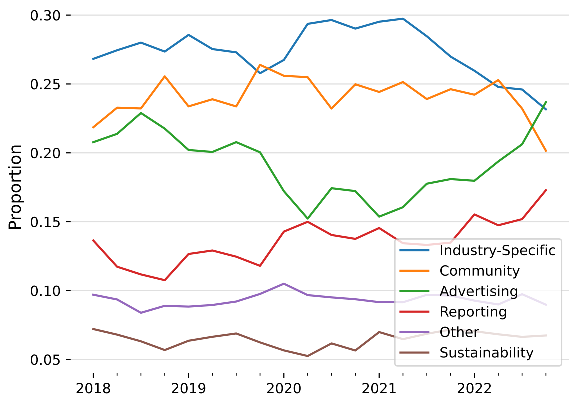

Figure 2: In 2018-2022, trends in the relative proportion of tweet categories vary, but suggest a possible shift of emphasis in recent years to include more corporate reporting (R) in addition to substantial discussion of community (C), with a pandemic-era dip in more straightforward advertising (A).

To focus on language from company accounts, we filtered out replies and retweets that were not "quote tweets", i.e., tweets that comment on an existing tweet. Finally, we pruned our vocabulary to terms (delimited by whitespace or punctuation) used by at least two companies. Without this processing, company-specific hashtags and terms overwhelmed our model. This brought the resulting vocabulary size from ∼1.36M to ∼52.5K across ∼827K tweets. The final dataset contains 827,403 tweets from 525 distinct Twitter accounts.

4. Analysis

4.1. Topic Model

To understand themes in our data, we wanted to find an unsupervised representation of similarity in our documents. We explored different models to identify themes in our corpora, including LDA topic modeling (Blei et al., 2003) and Sentence-BERT (Reimers and Gurevych, 2019) tweet embeddings. From our initial analysis, we decided to focus our work on a Biterm Topic Model (BTM) (Yan et al., 2013), which replaces the use of term-document frequencies with word co-occurrences in a 15-word window to respond better to shorter texts than LDA. The model outputs topics, or probability distributions over our vocabulary, with terms being allowed to have nonzero probability across multiple topics, and can be used to represent tweets as mixtures of topics. We trained our model on the full corpus of ∼827K tweets with 50 topics.

4.2. Topics

For each topic in our model, we inspected the 50 words of highest probability, the 10 tweets with the highest proportion of that topic, and top 10 Twitter accounts (by proportion of the account's tweets which were over a threshold for that topic). To develop themes across the topics, three authors manually labeled each topic, meeting in-person to resolve disputes on individual labels. Since the topic model captures all tweets in this period, we expect some topics to be less coherent; however, the authors did their best to understand why these words may have been grouped together using sample tweets. After labeling was established, we inductively developed six high-level topic categories to group related topics: **Industry-Specific Speech**, **Advertising**, **Corporate News and Reporting**, **Community and People**, **Sustainability**, and **Other** (which includes unclear or incoherent topics). We summarize these topics in Table 2 in the appendix. We plot data from years where we observed at least 100K total tweets before filtering. In this time period (2018-2022), we observe that there is a growth in corporate- and socially-focused speech, as shown in Figure 2.

We verify the existence of expected CSR themes anticipated by Stanislavská et al. (2023) around the environment, including (i) Sustainability (Topic 33 and 36), (ii) Climate (Topic 8), and (iii) Waste (Topics 12 and 40). We also see that keywords alone can be somewhat confusing for analysis: both topics 8 and 40 contain the words "sustainable" or "sustainability" 3 times within their top 50 words, but the difference is in how they use the word. Topic 8 focuses more on company announcements related to their sustainability efforts and goals (e.g. top document 10: "See our sustainability goals and progress achieved: https://t.co/GUUZEhJA26", @PPG), while topic 40 focuses on information and promotion of healthy sustainable practices (e.g. top document 9: "What's the wastewater and recycling connection? #WorldWaterDay https://t.co/qT8JU24eJR", @amwater). While the top companies in Topic 8 focus on sustainability (including Trane Technologies, Sempra, and NextEra Energy), we find The Coca-Cola Company (@CocaColaCo) ranked 7th for the topic, a company with a documented history of both a strong public CSR strategy and a record of significant environmental and social harm (Karnani, 2014). Similar overlaps occur in vocabulary for discussions of energy: Topics 33 and 36 focus on clean energy, while Topic 41 is focusing on energy production and Topic 35 mentions energy in the context of powering electronics.

The *Community and People* category also includes both internally-focused speech on excellent workplaces (e.g. topic 9, which revolves around

Predictors of High Risk		Predictors of Low Risk	
Topic 41 (58.67)	Topic 3 (27.81)	Topic 12 (-28.87)	Topic 13 (-27.97)
new	proud	new	new
energy	support	make	industry
gas	employees	help	learn
million	communities	packaging	latest

(a) Total Risk

Predictors of High Risk		Predictors of Low Risk	
Topic 41 (101.01)	Topic 8 (44.72)	Topic 10 (-33.07)	Topic 24 (-31.77)
new	energy	health	culture
energy	sustainable	help	work
gas	climate	access	inclusive
million	global	care	diversity

(b) Environmental Risk

Predictors of High Risk		Predictors of Low Risk	
Topic 16 (70.53)	Topic 28 (45.86)	Topic 12 (-80.04)	Topic 13 (-66.94)
people	patients	new	new
help	help	make	industry
work	treatment	help	learn
world	disease	packaging	latest

(c) Social Risk

Predictors of High Risk		Predictors of Low Risk	
Topic 30 (68.05)	Topic 16 (39.93)	Topic 13 (-37.41)	Topic 12 (-32.37)
risk	people	new	new
help	help	industry	make
global	work	learn	help
companies	world	latest	packages

(d) Governance Risk

Table 1: The strongest predictors of risk scores. This includes the top two topics that are the best predictors of high risk and the top two topics that are the best predictors of low risk, along with their respective top words and (parenthetical) regression coefficients.

highlighting workplace recognition and achievement) and external-focused communication (e.g. topic 47, which highlights supporting, donating to, and volunteering work). Prior work by Pilgrim and Bohnet-Joschko (December 2022) highlights social themes in CSR as a particular focus in digital media in ways that echo our topics, including the categories of (i) employee relations (Topics 21 and 26), (ii) diversity and inclusion (Topic 24), (iii) local community engagement (Topics 3 and 47), and (iv) philanthropy (Topic 43). We also see less specific socially-oriented topics like Topic 37. With simpler terms including "new", "help," "customers," and later "world", Topic 37 is led by McDonald's, and then immediately followed by multiple defense contractors and energy companies including Raytheon, HII, Lockheed Martin, and General Dynamics. This language connection between companies in seemingly unrelated industries suggests a possible trend of broad tweets about "helping the world", perhaps distinguishing a public CSR posture from concrete action and investment.

4.3. ESG Correlation

To understand how the learned topics from the Twitter corpus reflect corporate actions, we test whether topic proportions are predictive of 2022 Environmental, Social, and Governance (ESG) scores. These scores quantitatively describe Sustainalytics' assessment of companies based on exposure and management approaches to ESG risks. While Berg et al. (2022) show that ESG scores can disagree between sources, they highlight Sustainalytics as having the highest average correlation across other ESG metrics considered in their study. A low combined ESG risk score (<20) indicates positive work done towards managing ESG risks, while a higher risk score (>30) indicates greater concern. We use both the combined ESG score for each firm and three separate scores for Environment, Social, and Governance. We rescale each of these scores to a 0-100 scale for clarity of comparison. We obtained these scores for 453 of our companies via Yahoo Finance.

We represent each company using a 50-dimensional vector, where the i^{th} element is the proportion of the company's tweets in topic i. We then used ridge regression and Leave-One-Out (LOO) cross validation to try to predict both combined ESG and separate E, S, and G scores for each company. When computing regression, the ESG scores were all scaled to be from 0-100, by multiplying the environmental risk scores by 2 and the social and governance risk scores by 4. While fit was strongest for the environmental risk scores ($R^2 = 0.5$, RMSE $= 7.8$), it was weaker for the other two components, social ($R^2 = 0.16$, RMSE $= 13.4$) and governance ($R^2 = 0.16$, RMSE $= 7.94$), as well as for total risk ($R^2 = 0.22$, RMSE $= 6.2$). When compared to a baseline of predicting the risk as the averaging risk scores across the sector in our data, we see that only environmental scores are better predicted by our topic model than by the baseline

($R^2 = 0.4$, RMSE $= 8.9$).

However, even with low correlation, we still can find some meaningful trends in some of our topic features. We used coefficients from the regression model to find which topics were most predictive of high risk scores (most positive coefficients), as well as which predicted lower risk scores (most negative). The top words in these topics for prediction on each E/S/G score and their respective regression coefficients are presented in Table 1. While these are the most extreme, many more topics were also significant; from a permutation test, we determined that coefficients above 0.5 or below -1 were unlikely to be a result of random variation.

Unsurprisingly, we found the strongest predictor of high total ESG risk score was Topic 41, related to gas and energy companies, with a coefficient of 58.67. However, the next three highest predictors were topics that highlighted community "support" and "help" in abstract terms (Topic 3, 16, and 47, with coefficients 27.81, 21.26, 19.26). In contrast, we found that topics that related to concrete sustainable development practices and financial transparency correlated to lower risk scores.

Following Topic 41, the highest predictors of environmental risk was Topic 8, which contained speech relating to sustainability efforts, including #sustainability. The fact that a sustainability-focused topic indicates higher, not lower, risk, suggests that topic 8 is actually capturing "greenwashing" by companies to combat concerns about their climate practices. The other highest predictors also overlapped with those of total risk (Topics 3, and 47). The topics that indicated low environmental risk, Topics 10 (-33.07), 24 (-31.77), 4 (-30.18), and 30 (-27.96), contained more concrete words relating to healthcare, diversity and inclusion, and transparency about company finances.

Surprisingly, one of the highest predictors of social risk was relating to medical treatment and disease, potentially pointing to the complexity of intersecting profit motives with life-saving interventions. Topics that predicted a lower social risk included discussion of sustainable development (Topic 12) and technology reporting (Topic 13), with words inviting information exchange like "discuss", "opportunities", "solution", and "learn".

Finally, high governance risk was predicted by Topic 30 (68.05), containing words about corporate financial risk and economic impacts, as well as topics about community recognition (Topics 16, 43). Conversely, topics about technology solutions and sustainability that contained explicit references to environmental issues ("reduce", "carbon", "clean", "air", "emissions", "renewable") indicated a lower corporate governance risk.

5. Conclusion

In our work so far, we have collected a large corpus of corporate speech across 10 years of Twitter accounts for S&P 500 companies and trained a topic model to find patterns of discussion around CSR-focused themes. We found signs of both genuine reporting on CSR action from companies and cheap talk. The less explicit CSR focused topics correlated with increased ESG risk, especially those related to environmental concerns. These findings suggest that firms might be using communications about CSR as marketing strategies without fully investing in sustainability. We hope in our continuing work to reason further about individual variation in company language and concreteness/vagueness over time, as well as to compare Twitter behavior with spending data to show what distinguishes messaging of firms that invest funds towards sustainability and social good.

6. Acknowledgments

We would like to thank Eloise Burtis and the Pomona ITS team for their computing and data support. We thank our ECONLP 2024 reviewers for their kind reviews.

7. Bibliographical References

Theo B. Araujo and Jana Kollat. 2018. Communicating effectively about CSR on Twitter: The power of engaging strategies and storytelling elements. *Internet Research*, 28:419–431.

Najah Attig and Abdlmutaleb Boshanna. 2023. ESG washing: when cheap talk is not cheap! *Available at SSRN: https://ssrn.com/abstract=4672511*.

Andrew Baker, David F Larcker, Charles McClure, Durgesh Saraph, and Edward M Watts. 2023. Diversity washing. *Chicago Booth Research Paper*.

Florian Berg, Julian F Kölbel, and Roberto Rigobon. 2022. Aggregate Confusion: The Divergence of ESG Ratings. *Review of Finance*, 26(6):1315–1344.

Steven Bird, Ewan Klein, and Edward Loper. 2009. *Natural language processing with Python: analyzing text with the natural language toolkit*. "O'Reilly Media, Inc.".

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent Dirichlet allocation. *J. Mach. Learn. Res.*, 3(null):993–1022.

Howard R. Bowen. 2013. *Social Responsibilities of the Businessman*. University of Iowa Press.

Costanza Conforti, Jakob Berndt, Mohammad Taher Pilehvar, Chryssi Giannitsarou, Flavio Toxvaerd, and Nigel Collier. 2022. Incorporating stock market signals for Twitter stance detection. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4074–4091.

Richard M Crowley, Wenli Huang, Hai Lu, and Wei Luo. 2019. Do firms manage their CSR reputation? Evidence from Twitter. *Singapore Management University School of Accountancy Research Collection*, pages 1–44.

Elisabet Garriga and Domènec Melé. 2004. Corporate social responsibility theories: Mapping the territory. *Journal of Business Ethics*, 53:51–71.

Thomas F Johnson and Matthew P Greenwell. 2022. Are companies using Twitter to greenwash and hide bad environmental performance? *Energy, Ecology and Environment*, 7(3):213–226.

Armand Joulin, Edouard Grave, Piotr Bojanowski, Matthijs Douze, Hérve Jégou, and Tomas Mikolov. 2016a. Fasttext.zip: Compressing text classification models. *arXiv preprint arXiv:1612.03651*.

Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2016b. Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.

Aneel Karnani. 2014. Corporate social responsibility does not avert the tragedy of the commons. case study: Coca-Cola India. *Economics, Management, and Financial Markets*, 9(3):11–23.

Michael L Kent and Maureen Taylor. 1998. Building dialogic relationships through the world wide web. *Public Relations Review*, 24(3):321–334.

Abagail McWilliams and Donald S. Siegel. 2000. Corporate social responsibility and financial performance: correlation or misspecification? *Strategic Management Journal*, 21:603–609.

Shintaro Okazaki, Kirk Plangger, Douglas West, and Héctor D. Menéndez. 2020. Exploring digital corporate social responsibility communications on Twitter. *Journal of Business Research*, 117:675–682.

Marc Orlitzky, Frank L. Schmidt, and Sara L. Rynes. 2003. Corporate social and financial performance: A meta-analysis. *Organization Studies*, 24:403 – 441.

Katharina Pilgrim and Sabine Bohnet-Joschko. December 2022. Corporate social responsibility on Twitter: A review of topics and digital communication strategies' success factors. *Sustainability*, 14(24).

Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. *arXiv preprint arXiv:1908.10084*.

Svetlana Rybalko and Trent Seltzer. 2010. Dialogic communication in 140 characters or less: How fortune 500 companies engage stakeholders using Twitter. *Public Relations Review*, 36(4):336–341.

Camilla Salvatore, Silvia Biffignandi, and Annamaria Bianchi. 2022. Corporate social responsibility activities through Twitter: From topic model analysis to indexes measuring communication characteristics. *Social Indicators Research*, 164(3):1217–1248.

Jianfeng Si, Arjun Mukherjee, Bing Liu, Qing Li, Huayi Li, and Xiaotie Deng. 2013. Exploiting topic based Twitter sentiment for stock prediction. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 24–29.

Jianfeng Si, Arjun Mukherjee, Bing Liu, Sinno Jialin Pan, Qing Li, and Huayi Li. 2014. Exploiting social relations and sentiment for stock prediction. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1139–1145.

Lucie Kvasniková Stanislavská, Ladislav Pilař, Xhesilda Vogli, Tomá Hlavsa, Katerina Kuralová, Abby Feenstra, Lucie Pilařová, Richard Hartman, and Joanna Rosak-Szyrocka. 2023. Global analysis of Twitter communication in corporate social responsibility area: sustainability, climate change, and waste management. *PeerJ Computer Science*, 9.

US SIF Foundation. 2020. Report on US sustainable and impact investing trends 2020. Technical report, US SIF Foundation.

Xiaohui Yan, Jiafeng Guo, Yanyan Lan, and Xueqi Cheng. 2013. A biterm topic model for short texts. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13, page 1445–1456, New York, NY, USA. Association for Computing Machinery.

A. Full Topic List

The table in the following pages summarizes the topics by high-probability words and prominent companies. We include both our find-grained labels and our broader categorization of these topics: textbfIndustry-Specific Speech (I), **Advertising** (A), **Corporate News and Reporting** (R), **Community and People** (C), **Sustainability** (S), and **Other** (O).

#	Label	Top Words	Top Companies
0 (I)	Healthcare Sector	help care health new support provide program people access make	@ShopSimon @Take2Interactiv @tuicruises @RealtyIncome @EXPD_Official
1 (I)	Biotech	clinical new research development patients help cell drug discuss learn	@CatalentPharma @Incyte @CRiverLabs @CorningLifeSci @moderna_tx
2 (O)	Social Media	new latest episode shares future video look trends social blog	@nielsen @expediamedia @CBOE @GoldmanSachs @TrimbleCorpNews
3 (C)	Community Support	proud support employees communities students local community commitment efforts help	@VentasREIT @comcast @DevonEnergy @HeyCisco @PPLCorp
4 (R)	Financial Reporting	market · global prices supply bond demand high rose economic	@SPGlobal @Prologis @ICE_Markets @TRowePrice @MoodysInvSvc
5 (C)	Community Support	help work employees make better technology improve business new people	@Paycom @ServiceNow @Ceridian @kroger @Paychex
6 (C)	Community Support	health people help mental impact support care safety work important	@ElevanceHealth @Cigna @ViatrisInc @Humana @Centene
7 (R)	News Reporting	global markets economic market impact new credit pandemic growth insurance	@TRowePrice @MoodysInvSvc @BlackRock @FTI_US @GoldmanSachs
8 (S)	Sustainability	energy sustainable climate global future sustainability emissions carbon commitment #sustainability	@Trane_Tech @mhkgreenworks @sempra @nexteraenergy @Edison_Energy
9 (C)	Recognition	proud named year recognized honored list celebrate 2021 years 100	@KeurigPepper @Omnicom @VentasREIT @nexteraenergy @DowNewsroom
10 (I)	Healthcare Sector	health help access care support healthcare improve provide resources program	@Centene @cvshealth @ElevanceHealth @UnitedHealthGrp @ViatrisInc
11 (O)	Other	✓ help need new right business information know make online	@KelloggsUS @AskAmex @InsidePMI @VERISIGN @AltriaNews
12 (S)	Products and Packaging	new make help packaging products like food - work team	@WestRock @packagingcorp @BallCorpHQ @Sealed_Air @IntlPaperCo
13 (R)	Technology Reporting	new industry learn latest technology digital trends help supply experts	@McKesson @Gartner_Inc @PTC @healthcare_abc @Applied4Tech
14 (A)	Positive Advertising	new holiday favorite season time like make just best perfect	@Ross_Stores @LambWeston @RealReddiWip @bathbodyworks @smuckers
15 (I)	Information Technology	new business digital help data learn technology latest customer financial	@FISglobal @Fiserv @Broadridge @QuickBooks @StateStreet
16 (C)	People	people help work world make women like support helping we're	@Meta @tuicruises @3M @Intuit @AbbottNews
17 (O)	Other	make food like help time people know way water new	@VlasicStork @KeurigPepper @OpenTable @smuckers @pizzahut
18 (I)	Information Technology	data business security digital help learn discuss key organizations leaders	@Fortinet @Gartner_Inc @Protiviti @DXCTechnology @Equinix
19 (O)	Other (Short Hashtags)	- | / + – new love favorite great like	@skyworksinc @DukesMeats @ChipotleTweets @newell_brands @InvitationHomes
20 (R)	Financial Reporting	financial growth results quarter earnings market 2021 year strong new	@WECEnergyGroup @RealtyIncome @FactSet @MarathonOil @mhkgreenworks
21 (C)	People	new look team looking forward learn experience great time -	@iTeroScanner @poolcorp @IFF @amphenol @AmericanAir
22 (I)	Information Technology	new solutions technology data learn design help software technologies digital	@NXP @L3HarrisTech @ANSYS @MicrochipTech @Qualcomm
23 (I)	Financial Sector	volume near options • contracts trading futures million day term	@MarketAxess @CBOE @CMEGroup @ICE_Markets @FactSet
24 (C)	Diversity and Inclusion	culture work inclusive diversity diverse employees inclusion commitment women create	@Intuit @VentasREIT @ADP @KeurigPepper @Accenture_US

25 (I)	Biotech	learn new help using - design webinar cell process booth	@mettlertoledo @BioRadLifeSci @CorningLifeSci @BioRadFlowAbs @WatersCorp
26 (C)	Recognition	support team employees honor help work service thank members military	@HCAhealthcare @sbasite @ONEOK @UHS_inc @genuinepartsco
27 (O)	Other (Quantities)	- new million customers products years 2 support provide –	@PACCARFinancial @MarriottBonvoy @CharterNewsroom @Prologis @skyworksinc
28 (I)	Surgery, Medicine	patients help treatment disease heart care people risk patient cancer	@zimmerbiomet @DaVita @IntuitiveSurg @Abiomed @Hologic
29 (A)	Advertising	new look latest series collection – features — featuring iconic	@RalphLauren @Delta @EsteeLauder @Silversea @CarnivalPLC
30 (R)	Financial Reporting	risk help global companies risks impact financial challenges health need	@mercer @MarshGlobal @MarshMcLennan @GuyCarpenter @BRINKNewsNow
31 (C)	Power Service	power customers help stay weather safety crews safe outages report	@DominionEnergy @PSEGdelivers @EversourceMA @DTE_Energy @DukeEnergy
32 (O)	Other (Informal)	time tips know make help just sure you're need home	@Invisalign @OurTimeDating @hinge @KelloggsUS @Discover
33 (S)	Sustainable Energy	new team power energy help - future solar water make	@Enphase @nscorp @CSX @SolarEdgePV @CrownCastle
34 (A)	Events	learn today live event discuss - virtual join booth miss	@AristaNetworks @FactSet @ONEOK @LiveNation @IntuitiveSurg
35 (I)	Technology	power help new solutions make energy electric technology safety learn	@LKQCorp @autozone @IRProducts @ParkerHannifin @monolithicpower
36 (S)	Energy Responsibility	energy new help emissions reduce gas power carbon save electric	@Enphase @SolarEdgePV @Humana @EversourceMA @DTE_Energy
37 (A)	Advertising	new help customers look support business meet team make world	@McDonalds @RaytheonTech @WeAreHII @LockheedMartin @DukeEnergy
38 (A)	Advertising	win - chance time day booth sure ready new just	@exocad @UPS @AmericanAir @SlimJim @MonsterEnergy
39 (I)	Technology	power data new energy help performance solution customers network solutions	@monolithicpower @TXInstruments @MicrochipTech @Equinix @SEAGATE
40 (S)	Energy and Waste	water energy help gas reduce natural use waste air clean	@Pentair @RepublicService @amwater @Xylem @AOSmithHotWater
41 (R)	Energy Reporting	new energy gas million years largest - announced facility production	@Lindeplc @conocophillips @KeurigPepper @Kinder_Morgan @northropgrumman
42 (I)	Home Renovation	home like new space tips make kitchen room perfect living	@Lennar @PulteHomes @DRHorton @HomeDepot @LarsonDoors
43 (C)	Recognition, Announcements	proud excited team new announce support work students share sponsor	@TruistNews @tuicruises @genuinepartsco @Allstate @CaesarsEnt
44 (C)	Workplace	help work employees new career people business make need talent	@mercer @roberthalf @Paycom @Paychex @CamdenLiving
45 (I)	Healthcare	new data learn drug help using testing use development clinical	@thermofisher @WestPharma @WatersCorp @CatalentPharma @PerkinElmer
46 (R)	News, Announcements	new - people shares know latest like help learn —	@HeyCisco @travelocity @kroger @Orbitz @HLCruises
47 (C)	Community Support	help food support employees local team families communities million donated	@molinahealth @ConagraBrands @KelloggCompany @IDEXCorp @IntlPaperCo
48 (R)	Financial Reporting	latest new report year - 2021 credit 10 – impact	@VERISIGN @turbotax @creditkarma @TheHartford @CFIndustries
49 (A)	Time, Dating	. time years day team love @ summer #dating	@united @SherwinWilliams @kroger @OurTimeDating @Match

Table 2: The top 10 words and top 5 accounts for each topic. Each topic is hand-labeled with an approximate subject for the topic. Top words that include non-visible ASCII characters have been omitted, and the first 10 words with visible characters are included.