USER-GENERATED TWEETS ABOUT GLOBAL GREEN BRANDS: A SENTIMENT ANALYSIS APPROACH

KORISNIČKI GENERIRANE TWEET PORUKE O GLOBALNIM ZELENIM MARKAMA: PRISTUP ANALIZE SENTIMENTA

Saba Resnik[a], Mateja Kos Koklič[b]

[a] University of Ljubljana, Faculty of Economics, Kardeljeva ploščad 17, 1000 Ljubljana, SLOVENIA, e-mail: saba.resnik@gmail.com

[b] University of Ljubljana, Faculty of Economics, Kardeljeva ploščad 17, 1000 Ljubljana, SLOVENIA, e-mail: mateja.kos@ef.uni-lj.si

Abstract

Purpose – Microblogging platforms are generating an infinite volume of content on various topics. Therefore, traditional marketing methods can hardly be employed for its effective research, but sentiment analysis has recently emerged to cope with this challenge. While considerable academic effort has been devoted to investigating consumer behavior towards green brands, studies explicitly addressing consumer sentiments regarding such brands are still rare. Hence, we apply the sentiment analysis approach to investigate consumer sentiments towards 26 global green brands.

Design/Methodology/Approach – First, we collected a random set of user-generated tweets in English that were posted in a six-month period and included at least one of the selected global green brands. When classifying the posts, we extracted polarity information from a passage, resulting in values ranging from positive to negative.

Findings and implications – Based on a relative frequency word count, we found that consumers often express their sentiments about products, their characteristics and personal consequences of using them. Next, we analyzed average positive and negative consumer sentiments. As previously demonstrated, most tweets are either not strongly affective or they are ambiguous. Based on such

Sažetak

Svrha – Platforme za mikroblogiranje generiraju bezbroj sadržaja o raznim temama. Zbog toga se tradicionalne marketinške metode teško mogu koristiti, a nedavno se pojavila analiza sentimenta kako bi se borila s ovim izazovom. Unatoč znatnim akademskim naporima posvećenim potrošačevu ponašanju usmjerenom prema zelenim markama, istraživanja koja se eksplicitno bave potrošačevim mišljenjima o tim markama i dalje su vrlo rijetka. Stoga, primjenjujemo pristup analize sentimenta da bismo istražili mišljenja potrošača o 26 globalnih zelenih maraka.

Metodološki pristup – Prvo, prikupili smo slučajan skup korisnički generiranih tweet poruka na engleskom jeziku koje su objavljene u razdoblju od šest mjeseci i uključivale barem jednu od odabranih globalnih zelenih maraka. Pri razvrstavanju objava izvučeni su podatci o polaritetu iz odlomka, što je rezultiralo vrijednostima u rasponu od pozitivnih do negativnih.

Rezultati i implikacije – Na temelju relativne učestalosti broja riječi utvrdili smo da potrošači često izražavaju mišljenja o proizvodima, njihovim karakteristikama i osobnim posljedicama njihova korištenja. Nadalje, analizirali smo prosječna pozitivna i negativna mišljenja potrošača. Kao što je prethodno prikazano, većina tweet

MARKET
TRŽIŠTE

empirical insights, companies can better manage their brand perception on Twitter and other social media, as an integral part of their proactive marketing strategy.

Limitation – The study also has some limitations. It has a limited ability to reveal consumer motivations. The lexicon-based method we used may sometimes fail to recognize subtle forms of linguistic expression.

Originality – This research builds onto prior studies on green brands by applying sentiment analysis. It adds to the existing knowledge by investigating consumer sentiments towards 26 global green brands.

Keywords – microblogging, tweets, consumer sentiment, sentiment analysis, global green brands

poruka nije jako afektivna ili su dvosmislene. Na temelju takvih empirijskih uvida, poduzeća mogu bolje upravljati percepcijom svoje marke na Twitteru i drugim društvenim medijima, što bi trebalo biti sastavni dio njihove proaktivne marketinške strategije.

Ograničenja – Istraživanje ima nekih ograničenja. Ima ograničenu mogućnost otkrivanja motivacija potrošača. Korištena metoda bazirana na rječniku ponekad ne može prepoznati suptilne oblike jezičnog izraza.

Doprinos – Ovo istraživanje proširuje prethodna istraživanja o zelenim markama primjenom analize sentimenta. Dopunjuje postojeća znanja istraživanjem mišljenja potrošača o 26 globalnih zelenih maraka.

Ključne riječi – mikroblogiranje, tweet poruke, mišljenja potrošača, analiza sentimenta, globalne zelene marke

1. INTRODUCTION

The growing popularity of Web 2.0 is not only strongly aligned with consumers' passive absorption of online content but also with their creation, distribution, and exploitation of online content. A shift that has been noted in the past few years is the rise of user-generated content (in contrast to firm-created content) (Ceballos, Crespo & Cousté, 2016). Hence, today's challenge is to retrieve relevant data and transform it into actionable information (Montoyo, Martiniz-Barco & Balahur, 2012).

The exponential growth in evaluative data-rich resources associated with Web 2.0, such as online forums, web blogs, and microblogging services, has generated a huge universe of content rich in public opinion on a wide array of subjects (Gunter, Koteyko & Atanasova, 2014). Opinions expressed on social networks as the most prominent Web 2.0 platforms significantly influence public behavior across diverse areas, including the purchase of products and services or the shaping of political views (Eirinaki, Pisal & Singh, 2012). Since the volume of such content is infinite, especially on microblogging services, a traditional content analysis can hardly be employed for its effective research (Okazaki, Diaz-Martin, Rozano & Menendez-Benito, 2014).

Consequently, several studies have recently applied sentiment analysis (SA) as a suitable tool for coping with large amounts of marketing data in order to investigate brand perception, brand loyalty, and brand advocacy (e.g. Hu, Bhargava, Fuhrmann, Ellinger & Spasojevic, 2017). Sentiment analysis aims to examine consumer sentiments defined as tacit, context-specific explanations of consumer feelings, experiences, and emotions about a product or service (Hu, Koh & Reddy, 2014). The importance of consumer sentiments has been further underlined by the finding that they have been recognized as early indicators of consumer attitudes – one of the cornerstones of consumer behavior (O'Connor, Balasubramanyan, Routledge & Smith, 2010). Hence, uncovering consumer sentiments might be of strategic importance to companies across various industries. Indeed, sentiment monitoring might enable companies to improve product quality and services, assess the impact of promotional campaigns, drive sales, and identify new business opportunities. Another highly relevant outcome of measuring sentiment is gauging users' perceptions of companies (Hu et al., 2017; Jansen, Zhang, Sobel & Chowdury, 2009). Given their pressing environmental and sustainability-related concerns, companies have been carefully developing their identities as "green brands". Their aim, in doing so, is to appeal to consumers using environmental responsibility as an important element of their competitive advantage (Wang, 2017).

This study attempts to explore consumer sentiments about major global green brands as expressed via Twitter by applying sentiment analysis. The contribution of this study is threefold. First, the most relevant contribution lies in exploring spontaneous expressions of consumer sentiment towards several global green brands. Various scholars have called for a deeper examination of consumer perceptions of green brands to shed more light on the persistently demonstrated gap between consumer perceptions of a company's greenness and its actual sustainability performance (Cordeiro & Seo, 2014; Interbrand, 2014). Hence, cognizance of consumer attitudes and sentiments can substantially improve the understanding of this gap. In addition, studies addressing consumer perceptions of green brands as communicated via microblogs have been rather sporadic, despite the relevance of green issues (Hoepner, Dimatteo, Schaul, Yu & Musolesi, 2017). Second, our study provides an insight into the relationships between brand dispersion and performance measures, using the contemporary approach of sentiment analysis and a realistic user-generated dataset. In particular, recent studies have indicated that brand dispersion, defined as variance in brand ratings across consumers, leads to reduced performance and reduced firm risk (Luo, Raithel & Wiles, 2013). Brand dispersion re-

flects brand polarization and can be captured by applying sentiment analysis. Also, our study complements previous research of Luo and others (2013) by providing a set of data obtained through a completely different method. Third, the present study is welcome due to a dearth of literature that employs opinion mining techniques – more specifically, sentiment analysis (e.g. Jansen et al., 2009; Mostafa, 2013). Cai, Spangler, Chen and Zhang (2010) emphasized the importance of focusing on sentiment monitoring; they argued that the "voice of the web" is important for revealing consumer, brand, and market insights.

The remainder of the paper is organized as follows. First, we provide the background to Twitter as a microblogging platform. Then, we explain the concept of sentiment analysis and present the role of marketing research on green brands. Next, we describe the methodology in detail and provide exploratory and quantitative analyses of the gathered data. Finally, we conclude with a discussion of the findings, limitations, and future research directions.

2. THEORETICAL BACKGROUND

The emergence of Web 2.0 has drastically altered the way users perceive the Internet by improving information sharing and collaboration (Kontopoulos, Berberidis, Dergiades & Bassiliades, 2013). The need to express a particular point of view and feelings about specific topics is inherent in human nature. Social media, as one of the cornerstones of Web 2.0, have opened up new possibilities for people to interact and express themselves (Bravo-Marquez, Mendoza & Poblete, 2013).

Social media can be described as a two-way communication platform that allows reciprocal communication between companies and users, and users and users (Liu & Shrum, 2002). It has both driven and coincided with a dramatic change in the way of communication. Du and others (2015) illuminate the potential

of using social media opinion mining research as a promising alternative to survey and polling for researchers and practitioners alike. An increasingly established category within social media is microblogging, which in essence encompasses broadcasting of brief messages to some or all members of the sender's social network through a specific web-based service (Kaplan & Haenlein, 2011). The need to consider microblogging in today's competitive world is substantiated by the fact that numerous microblogs mention a brand name. Interestingly, 80 % of Twitter users mentioned a brand in their tweets (Orban, Nagy, Kjarval & de Carmona, 2014). Microblogging has evolved into a practical way of sharing opinions on almost all aspects of everyday life (Kontopoulos et al., 2013). Consumers use microblogging to inform others of what they are doing or thinking, to obtain information, to share information, and to forward news and articles (Wood & Burkhalter, 2014). In the cases where a certain brand is mentioned, users might comment on products, services, and events held by the company, or else respond to the company's promotions (Jansen et al., 2009). Compared to traditional blogs, microblogs are strictly constrained in content size, but still enable users to post their opinions, experiences, and queries on various topics (Kaplan & Haenlein, 2011).

One of the most widespread online microblogging services is Twitter. It enables its users to send and receive posts, known as "tweets", consisting of up to 140 characters. This character limitation results in users being more concise and eventually more expressive than via other social networks and blogs. Additionally, tweets can be processed more effectively compared to lengthy blogs or articles (Kontopoulos et al., 2013). Tweets might contain different forms of content, such as images, text, videos, and interactive links (Twitter, 2018). According to Young (2010), Twitter has significantly lowered the barriers to creating content, which is why users easily share their day-to-day lives. Besides the restricted length, other features of Twitter messages are the casual language style, mixed

use of symbols and words, and high frequency of grammar and spelling errors (Du et al., 2015). Twitter currently records 330 million monthly active users who send over 500 million tweets per day (Aslam, 2018), mostly written in the English language (Mocanu et al., 2013). One of the characteristics of Twitter is also a varied pool of authors; therefore, it is possible to collect posts of users from different social and interest groups (Pak & Paroubek, 2010).

In contrast to traditional consumer surveys, consumers normally post their opinions on microblogging platforms without any external trigger or specification of topic (Schindler & Decker, 2013). The fact that these opinions are highly unlikely to be biased, they display a high level of authenticity, and are affective in their nature makes them appealing to the majority of readers. As a result, the analysis of freely expressed customer opinions and related concepts, such as sentiments and evaluations, is a promising alternative to conventional survey techniques (Decker & Trusov, 2010).

Sentiment analysis is described as the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities, such as products, services, organizations, individuals, issues, events, topics, and their attributes (Liu, 2012). In fact, rather than answering surveys about products and services, consumers freely express their thoughts and emotions on social media (Hu et al., 2017). Sentiment analysis tasks can be done at several levels, as suggested by Kumar and Sebastian (2012): word level, phrase or sentence level, document level, and feature level. Techniques for sentiment analysis can be broadly categorized into two classes of approaches. At the word level, sentiment analysis methods fall into the following two categories: (1) dictionary-based approaches (Kumar & Sebastian, 2012) and (2) machine learning approaches (Pang & Lee, 2008).

Alternatively, sentiment analysis can be described as a technique for identifying the ways in which sentiments are expressed in the text.

It can also be used to determine whether the sentiments represent positive or negative feelings about a specific product or service (Nasukawa & Yi, 2003). In fact, the majority of studies examine the polarity of expressed sentiments. For instance, Ortigosa, Martín and Carro (2014) implement this method by further refining its protocol to extract the sentiment polarity and detect significant emotional changes in Facebook messages. Previous research on sentiment analysis has also included evaluations of product reviews (Fang & Zhan, 2015; Kang, Yoo & Han, 2012), sentiment analysis of online news articles and feeds (Moreo, Romero, Castro & Zurita, 2012), as well as online forums and discussion boards (Abbasi, Chen & Salem, 2008; Homburg, Ehm & Artz, 2015). In the field of politics, researchers have used sentiment analysis to determine the sentiments expressed in tweets (Hu et al., 2017; Wang, Can, Kazemzadeh, Bar & Narayanan, 2012). Additionally, researchers have employed sentiment analysis in the field of tourism (Alaei, Becken & Stantic, 2017; González-Rodríguez, Martínez-Torres & Toral, 2016) and for differentiating between informative and emotional social media content (Denecke & Nejdi, 2009).

As evidenced, sentiment analysis has been applied in various fields of study. In addition, since sentiment analysis measures the polarity of brands, it can be utilized to examine brand dispersion – a metric with a significant impact on firm value and stock market performance. Luo and others (2013) demonstrated that brand dispersion is not only consistently related to lower abnormal returns, but is also a beneficial reduction in idiosyncratic risk. Moreover, their study showed that downward dispersion is more closely associated with returns than upside dispersion is.

Sentiment analysis on microblogs is receiving more and more attention from scholars since it contains important information stating either positive or negative feelings in a very limited space (Chamlertwat, Bhattarakosol, Rungkasiri & Haruechaiyasak, 2012). It can be used to moni-

tor Twitter in real time in order to detect major events and users' reactions to events (Benhardus & Kalita, 2013; Hsieh, Moghbel, Fang & Cho, 2013). Sentiment analysis research also includes the impact of tweets on movie sales (Rui, Liu & Whinston, 2013), public mood and emotion analysis (Bollen, Pepe & Mao, 2011), detecting irony (Reyes, Rosso & Buscaldi, 2012), specifying users' sentiments regarding different topics, and analyzing brand-related tweets (Ghiassi, Zimbra & Lee, 2016; Ghiassi, Skinner & Zimbra, 2013). Twitter offers a unique dataset in the world of brand sentiment since brands receive sentiment messages directly from consumers in real time in a public forum. Both the targeted and the competing brands have the opportunity to dissect these messages to determine potential changes in consumer sentiment. Taking advantage of these messages, however, requires researchers to deal with analyzing an immense amount of data produced by Twitter users each day (Ghiassi, Skinner & Zimbra, 2013).

Despite the wide range of topics analyzed with the help of sentiment analysis, only a handful of studies have probed into sentiment analysis on sustainability-related topics, as an area of immense relevance to various stakeholders. For example, Du and others (2015) illuminated the potential of using social media opinion-mining research as a promising alternative to surveys and polling for both researchers and practitioners. In the last few decades, consumer concern for restoring the ecological balance and a significant increase in the presence of green brands have been observed in the marketplace (Chen & Chai, 2010). Green brands have been defined as a set of attributes and benefits relating to the reduced environmental impact of these brands and their perception as being environmentally sound (Hartmann, Apaolaza Ibáñez & Forcada Sainz, 2005). Consumers tend to place importance on the environmental side of sustainability (Hanss & Böhm, 2012; Hosta & Žabkar, 2016).

Given their growing importance, scholars have also become more interested in investigating green consumer behavior, consumer attitudes, and sentiments regarding green brands. As a result, the scope of academic research has mirrored the relevance of environmental sustainability. Studies on green consumers encompass a wide range of topics: from the socio-demographic and psychographic profiling of green or environmentally conscious consumers (e.g. Akehurst, Afonso & Martins Gonçalves, 2012; Roman, Bostan, Manolică & Mitrica, 2015) to investigating drivers of, and barriers to, green consumer behavior (e.g. Minton, Kahle & Kim, 2015; Papista, Chrysochou, Krystallis & Dimitriadis, 2017; Tan, Johnstone & Yang, 2016). Accordingly, extant research has addressed individual-level determinants, such as the consumer's personality and socio-demographic characteristics (e.g. Lu, Chang & Chang, 2015) as well as the attitudes towards specific green products, green behavior, or green marketing in general (e.g. Borin, Lindsey-Mullikin & Krishnan, 2013; Olsen, Slotegraaf & Chandukala, 2014). The existing literature also offers insight into the role of values, beliefs, and norms in guiding consumers' pro-environmental intentions and behaviors (e.g. Han, 2015; Steg, Bolderdijk, Keizer & Perlaviciute, 2014).

Consumers' favorable attitudes towards green products and practices are often inconsistent with their actual behavior. Hence, a substantial body of literature on green consumer behavior probes into this widely acknowledged attitude-behavior or intention-behavior gap (e.g. Aschemann-Witzel & Niebuhr Aagaard, 2014; Biswas & Roy, 2015; Johnstone & Tan, 2015; Moser, 2015). These studies address why consumers' favorable attitudes towards green products and practices are often inconsistent with their actual behavior.

Additionally, companies are investing considerable efforts into associating environmental issues with brands and emphasizing the importance of environmental sustainability (Rios, Martinez & Molina, 2008). However, consumers do not always recognize a company's green efforts, as demonstrated by Cordeiro and Seo (2014), who found a significant gap between consumer

green brand recognition and companies' actual environmental performance. Further, Interbrand (2014) identified several global green brands where the brand's environmental performance was either significantly higher or significantly lower than consumer perceptions of that performance. For example, Interbrand identified Cisco, Nokia, and L'Oréal as brands with a better actual performance than what consumers perceived. Conversely, McDonald's, Coca-Cola, and Disney were perceived to be more sustainable than they actually were. Both cases underscore the importance of aligning consumer knowledge with the company's initiatives (Cordeiro & Seo, 2014).

3. METHODOLOGY

This study employs a sentiment analysis approach to investigate consumer sentiments regarding a selection of global green brands. In exploring these sentiments, we applied a five-step procedure suggested by Rambocas and Gama (2013): (1) data collection; (2) text preparation; (3) sentiment detection; (4) sentiment classification; and (5) presentation of output (Figure 1).

tage of doing so is the objectivity of the evaluative criteria used by Interbrand, which has been well established for many years and employed in several research studies (e.g., Chehab, Liu & Xiao, 2016; Wang, 2016). The initial list of 50 brands spanned 11 countries and 13 industries and was based on consumer perceptions as well as company performance – from corporate governance to management commitment, through the supply chain and, ultimately, the product and/or service. Although those brands were spread across major industries, our focus was the category of fast-moving consumer goods as a category that consumers are well acquainted with and are closely related to in their daily life. Consequently, we assumed that these items would most likely be the subject of tweeting. The initial list was reduced to 34 global green brands, including: Adidas, Apple, AXA, Canon, Cisco, Coca-Cola, Colgate, Danone, Dell, Disney, General Electric, Heineken, H&M, IBM, IKEA, Intel, Johnson & Johnson, Kellogg's, L'Oreal, McDonald's, Microsoft, Nestlé, Nike, Nokia, Panasonic, Pepsi, Philips, Samsung, Santander, Siemens, Sony, Starbucks, Xerox, and ZARA. This yielded a total of 133,178 posts mentioning the brands under scrutiny.

FIGURE 1: The process of sentiment analysis
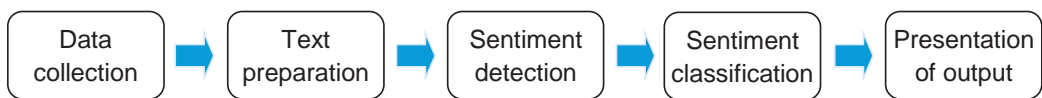
In stage 1, we collected a random set of user-generated tweets in English posted between 1 September 2014 and 10 March 2015 that included at least one of the selected green brand names. More specifically, since only the content expressed in words was collected for this study, other types of content, such as images and videos, were not included in our analysis. In case the text incorporated interactive links, they were also included. In order to identify relevant green brands, we relied on the Best Global Green Brands 2014 survey conducted by the brand consultancy Interbrand (2014). The main advan-

In the next state, we screened the extracted data to identify content irrelevant to the area of our study. In doing so, we excluded brands with less than 100 posts, namely the following: AXA, Danone, General Electric, H&M, Johnson & Johnson, Kellogg's, Santander, and Xerox. This resulted in the final data set containing 133,029 random tweets in the English language mentioning any of the 26 global green brands.

When classifying posts, a researcher might focus on various aspects of opinions: polarity, emotions, or strength (Hu et al., 2017). Our fo-

cus was on the polarity with the aim to extract polarity information from a passage, resulting in values ranging from positive to negative. Several methods can be used for identifying the orientation of sentiment (Miao, Li & Zeng, 2010). We used the lexicon-based method that requires a pre-defined dictionary of words, WordNet, which is commonly used for assessing positive or negative sentiments. WordNet is a large lexical database that contains nouns, verbs, adjectives, and adverbs, which are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. The main relation among the words in WordNet is synonymy, where words denote the same concept and are interchangeable in many contexts. WordNet labels the semantic relations among words, where the most frequent relation among synsets is the super-subordinate relation, which links more general synsets like {furniture, piece_of_furniture} to increasingly specific ones like {bed} and {bunkbed}. WordNet also labels relations such as meronymy semantic relation, verb synsets arranged into hierarchies, adjectives organized in terms of antonymy, and relations across parts of speech (POS; cross-POS relations) (Fellbaum, 2005; Princeton University, 2010).

For the purposes of this study, we used a simple module pattern.en combined with WordNet. This combination bundles a lexicon of adjectives (e.g. good, bad, amazing, irritating) that occur frequently in product reviews, annotated with scores for sentiment polarity (positive ↔ negative). The sentiment is determined by comparing tweets against the expert-defined entry in the dictionary, making it easy to determine the polarity of a specific sentence. For example, words that express a desirable feeling, such as "great" or "excellence", have a positive polarity while words that express an undesirable feeling, such as "bad" or "awful", have a negative polarity. Based on the adjectives it contains, the sentiment function returns the polarity value for the given sentence between -1.0 and +1.0, and

we used these scores to determine whether a single tweet post is positive, neutral, or negative (Liu, 2015). More specifically, we employed mathematical optimization to determine three segments: positive (from 0.19554 to 1.0), neutral (from 0.19028 to 0.19553), and negative (from -0.19029 to -1.0). In case the posts could not be defined as positive or negative, that is, if they were not expressing a feeling or an opinion, the sentiment analysis gave a neutral evaluation (Davis & O'Flaherty, 2012). For example, in our study, the sentiment function returned the polarity value +0.78 for the tweet post 'Fucking love Ikea!!' ({'polarity': 0.78125, 'text': 'Fucking love Ikea!!'}); therefore, we defined this tweet as positive, while defining the value -0.3 for the tweet post '@WallBlume_ eww i never liked Heineken' ({'polarity': −0.3, 'text': '@WallBlume_ eww i never liked Heineken'}) as negative.

The final step in the process of analyzing sentiments is to present the output, which is outlined in more detail in the Analysis of the data section. For the purpose of pre-testing our sentiment analysis tool, we also conducted a short online survey with 27 respondents who evaluated the polarity of 63 randomly selected tweets as positive, neutral, or negative. Their assessments matched the sentiment score produced by the automated tool in 71 % of cases. This overall agreement level is relatively close to the outcome of previous studies; for example, Wilson, Wiebe, and Hoffmann (2005) found an 82 %-match between annotators' judgments and automated judgments. We estimated that the level was sufficient to proceed with the analysis.

4. DATA ANALYSIS

4.1. Exploratory data analysis

In the first phase of our empirical analysis, we conducted exploratory research by generating relative frequency word counts. This step provides a researcher with insight into a particular topic or even predicts characteristics of the topic analyzed (O'Leary, 2011; Mostafa, 2013).

Namely, the frequency of appearance of terms is a widely used measure of interest of a specific topic, as shown by Asur and Huberman (2010); they found that the number of appearances of discussion about a single topic can be used to predict characteristics of the topic. Before deciphering the frequency of individual words, we looked at how many posts were associated with the selected brands. We found that the leading brand is Apple with 46,612 tweets, followed by Disney with 19,414 tweets, and Nike with 14,788 tweets. Figure 2 illustrates more specifically the number of tweets for those brands that were mentioned at least 1,000 times.

FIGURE 2: Total number of user-generated tweets for individual brands
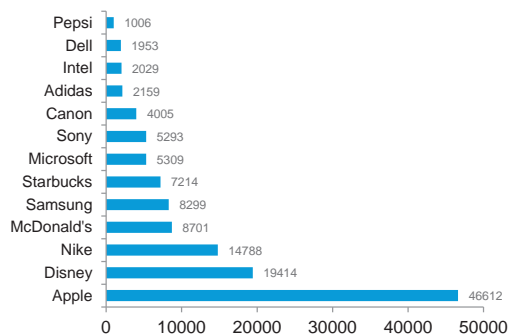
Note: * Brands mentioned fewer than 1,000 times are not displayed

Subsequently, we analyzed high-frequency words in tweet posts using the Hermetic Word Frequency Counter 15.52 software. If focusing on the Apple brand as that with the highest number of tweets, we noticed that by far the most commonly occurring word in the tweets is "watch" and this might be explained by the hype surrounding the introduction of the new Apple Watch to the market. Along these lines, the words "new", "iphone", "macbook", and "reveals" also reflect consumers' close monitoring of the company's actions (Table 1). In relation to the second most frequently mentioned brand – Disney, the following words had the highest

frequency: "movie", "life" or "live", "world", "dumbo", and "channel" (Table 1). Among the top 10 words, we also identified "old", "burton", and "tim", corresponding to an announcement that the American film director Tim Burton will direct a live-action remake of the Disney classic Dumbo (Fritz, 2015). Finally, we also investigated the most frequent words associated with Nike. As anticipated, the most commonly mentioned word is "air" related to the Nike Air technology and Max Air Technology. Other product-related terms among the top mentions were "size", "jordan", "max", "shoes", and "retro". Given that footwear sales are the largest source of revenue for Nike (Iyer, 2015), this outcome is somewhat expected. In addition, the word "just" appeared in slightly less than 12 % of tweets containing the brand name "Nike", presumably in most mentions referring to the Nike logo "Just Do It". The top 10 words also included the words "nikeplus" and "ran", presumably reflecting the growing popularity of running (Scheerder, 2015).

For illustration purposes, we also provide word frequency data for some of the selected green brands. Words such as "nba" and "shoes" have the highest frequency for Adidas, while words like "table", "furniture", and "gift" have the highest frequency for Ikea; words such as "served", "coffee", "awesome", and "morning" have the highest frequency for Starbucks, and words like "camera", "digital", "photography", and "kit" have the highest frequency for the brands Canon and Panasonic. As expected, we also found that competitive brands are frequently mentioned together in a single tweet, such as Adidas and Nike, Samsung and Apple, Pepsi and Coca-Cola, Canon and Nikon. Following the previously established notion that the majority of tweets are neutral or ambiguous (Mostafa, 2013), we additionally re-examined the top 25 words associated with each of the 26 selected brands to identify strongly emotional articulations. Being aware of a serious limitation of such analysis, namely the absence of the context (words are analyzed in isolation), we could only characterize a few emotionally intense words. Associated with the McDonald's brand and the Pepsi brand was the

word "fuck" (in 5.4 % and 16.9 % of tweets, respectively). Pepsi was also related to the words "wtf" and "dicks" (in 16.8 % and 16.7 %, respectively), while Starbucks was associated with positively charged words "awesome" (in 3.5 % of tweets) and "love" (in 2.9 % of tweets). However, in most cases the frequently mentioned words by themselves have no strong connotation.

lines the fact that the majority of the tweets fall into the neutral category by having a sentiment score between -0.2 and +0.2. In our study, the share of neutral tweets ranged from 71.56 % (ZARA) to 95.74 % (L'Oreal), indicating that a large percentage of tweets are not very affective. This is in line with the findings of other authors, for example, Lindgren (2012) and Mostafa

TABLE 1: Word frequency for tweets mentioning Apple, Disney, and Nike

APPLE			DISNEY			NIKE		
Word	Number	%	Word	Number	%	Word	Number	%
watch	20757	44.5	movie	1557	8.0	air	4824	32.6
new	2780	6.0	life	1269	6.5	size	1805	12.2
iphone	2665	5.7	world	1244	6.4	just	1750	11.8
macbook	2193	4.7	dumbo	1052	5.4	jordan	1679	11.4
reveals	2049	4.4	channel	918	4.7	max	1678	11.3
amp	2024	4.3	old	833	4.3	shoes	1579	10.7
edition	1778	3.8	burton	783	4.0	nikeplus	1265	8.6
dubai	1613	3.5	tim	772	4.0	new	1158	7.8
read	1448	3.1	throwback	765	3.9	ran	1087	7.4
now	1416	3.0	magic(al)	723	3.7	retro	1084	7.3

Note: The column displays the percentage of word frequency with respect to the total number of posts per brand.

Interestingly, when looking at the 25 most frequently mentioned words for any of the 26 brands, none of them is related to "green" or "sustainable" even though these brands were listed as the top global green brands. This indicates that, in general, consumers do not reflect on the companies' greenness when tweeting, but focus more on products, their characteristics, and personal consequences of using these products.

4.2. Overall sentiment scores

In the second part of the study, we conducted quantitative research to achieve the sentiment scores. Our aim was to examine the distribution of sentiment scores, which were divided into five categories: from -1 to -0.6, from -0.6 to -0.2, from -0.2 to +0.2, from +0.2 to +0.6, and from +0.6 to +1 (Figure 3). This visualization under-

(2013), who observed that it is very common for sentences in the analyzed dataset to be ambiguous or neutral and thus hard to place on either side of the continuum. As evident in Figure 3, the share of positive tweets exceeded the share of negative tweets for all brands except McDonald's, where the situation is reverse.

Next, we focused on the positive and the negative tweets to shed more light on potential asymmetry in consumer sentiments about the selected brands. More specifically, average positive and average negative sentiment scores were obtained for each brand (Figure 4). Brands such as Heineken (+0.43), Starbucks (+0.40), Colgate (+0.38), and Zara (+0.38) have the highest average positive score, while IBM (+0.26), Siemens (+0.27), and Philips (+0.28) are among the least positively evaluated brands. This partially corresponds to the sentiment distribution in

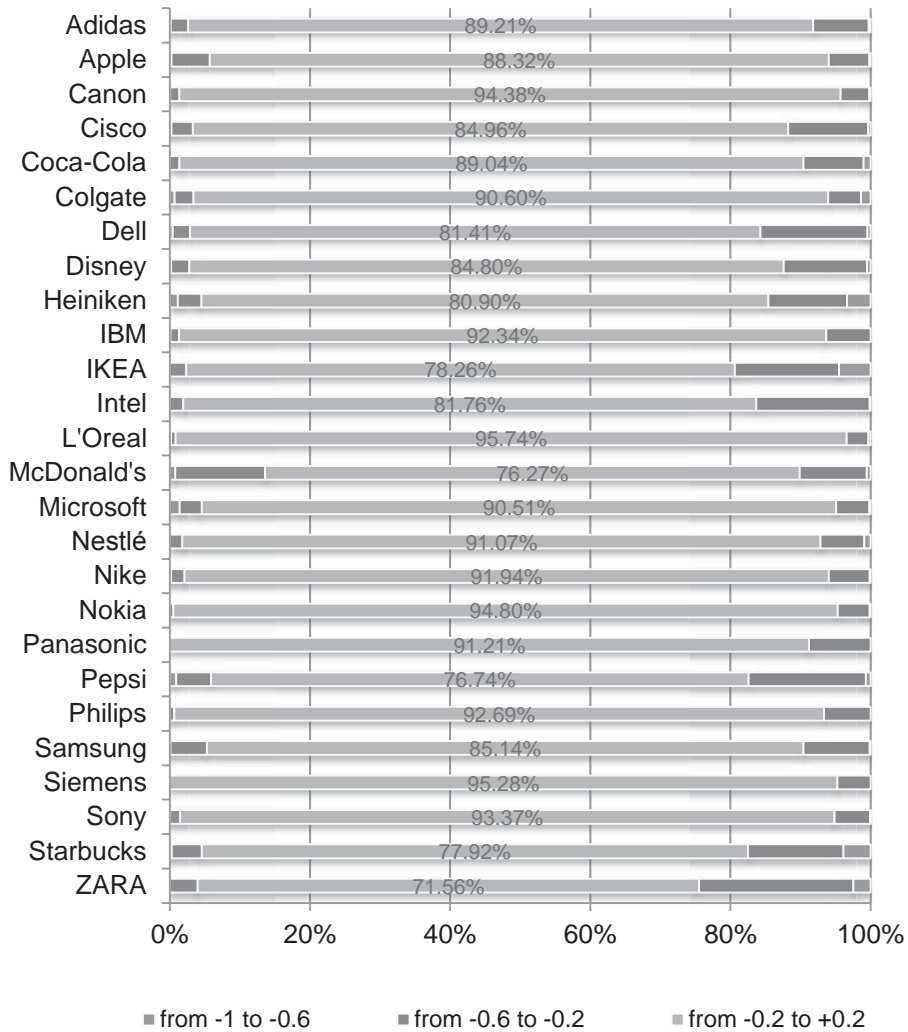FIGURE 3: Distribution of sentiment scores for green brands (% tweets)

Figure 3, showing that the IKEA, Starbucks, and Heineken brands received the highest percentage of maximally positive tweets (with scores from +0.6 to +1). On the other hand, the following brands have the lowest average negative scores: Nokia (-0.55), L'Oreal (-0.52), Heineken (-0.44), Dell (-0.43), and Microsoft (-0.42). Siemens had no negative tweets, therefore its negative score was 0.00 (the total number of tweets was low, i.e. 127), with Panasonic (-0.20) and Coca-Cola (-0.18) also receiving relatively low negative sentiment scores. This can be paralleled to the findings based on Figure 3 where Microsoft, Heineken, and McDonald's were assigned the highest share of very negative tweets (with scores ranging from -1 to -0.6). Some of these brands, such as Colgate (+0.38 and -0.42), Heineken (+0.43 and -0.44), and Zara (+0.38 and -0.41) tend to evoke strongly positive as well as strongly negative sentiments in consumers. In this case, consumers seem to be more divided in their affective stance on these brands.

FIGURE 4: Average negative and average positive sentiments toward green brands
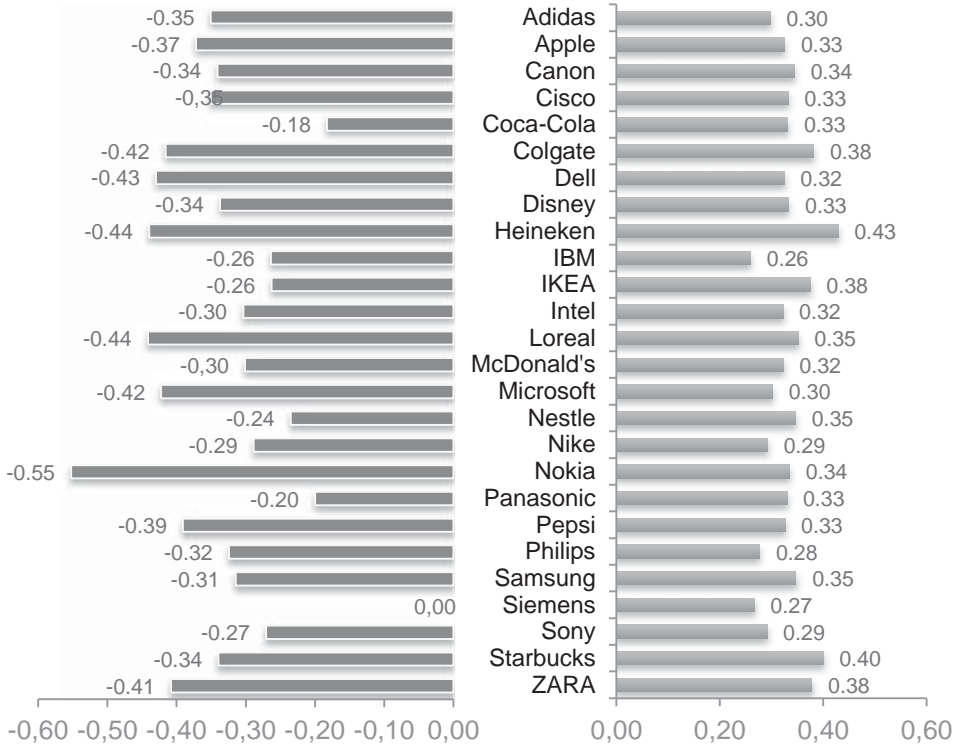
Finally, we provide a more realistic insight into the strength of the positive vs. negative scores by computing the total score per brand. When considering the overall average sentiment scores, all brands achieve positive values just slightly above zero, ranging from +0.1 to +0.11. The highest score was obtained by Philips and ZARA (in both cases +0.11), although the relative value of the average negative score is higher than the relative value of the average positive score. This indicates that the number of positively charged tweets was still higher, outweighing those negatively charged ones. Slightly lower but still noteworthy is the total score for the IKEA and Intel brands (+0.10). In these two cases, the average positive scores were higher than the average negative scores. In contrast, the least preferred brand according to the total score estimate is McDonald's with +0.01 as the average total score, followed by Apple and Microsoft both with a score of +0.03.

Following recent calls to measure brand dispersion as an indicator of potential brand inconsistency (Hu et al., 2017), we applied this brand metric to our dataset. Hence, we calculated the standard deviations of the sentiment scores of users' tweets, which were grouped according to their brand mention. As presented in the Appendix, brands with the lowest standard deviations (SD) are Siemens (SD = 0.803), Philips (SD = 0.842), Sony (0.896), IBM (SD = 0.897), and Canon (SD = 0.901). On the other side of the spectrum are brands with the highest deviations: McDonald's (SD = 1.885), Starbucks (SD = 1.904), Zara (SD = 1.904), Pepsi (SD = 2.007), and Heineken (SD = 2.107). Based on our summary table (provided in the Appendix), it may be concluded that brands with lower dispersion or polarization (lower SD) have a higher green rank. The majority of high-dispersion (high SD) brands tend to have a negative perception gap, meaning that consumers perceive these com-

panies to be more green or sustainable than they actually are. A few other observations can be made on the basis of the standard deviation and performance measures. When comparing five brands with the lowest dispersion and five brands with the highest dispersion, it could be stated that, on average, low-polarizing brands have a higher variation in stock price (difference between the highest and lowest prices during one year). Nonetheless, among the higher-dispersion brands, Starbucks also experienced a particularly high variation in stock price, both in 2014 and 2015. High-dispersion brands all recorded stock price growth in 2014 and 2015, except for McDonald's, which faced a slight drop in its stock price in 2014. In contrast, the majority of low-dispersion brands recorded a decrease in their stock price, as well as a decrease in their brand value in 2015. When comparing average revenues for top and bottom five brands (in terms of dispersion), it can be noted that the revenues of low-dispersion brands are almost twice as high as the revenues of high-dispersion brands.

5. GENERAL DISCUSSION

Our core focus in this study was to analyze consumer sentiments with regard to an assortment of 26 brands recognized by Interbrand as global green brands in 2014. We specifically focused on fast-moving consumer goods companies due to consumers' daily interaction with their products. Consequently, a more diverse and abundant landscape of consumer sentiments was expected. Indeed, the selected brands seem to differentiate on various levels. For example, notable differences were found in the number of brand mentions, with Apple being far ahead of other brands, supposedly as a result of being the most valuable company worldwide among other reasons. Apple was awarded this position based on its brand value, estimated as the likely future sales that are attributable to a brand and a royalty rate that would be charged for the use of the brand (Brand Finance, 2016).

When analyzing word frequency in the pool of tweets, we found that the words related to products and their outcomes are often among the most frequently mentioned. This finding is consistent with the notion of Du and others (2015) that tweets about sustainability are relatively rare compared to those discussing "hot" topics. Perhaps this pertains to the inherent nature of consumers who respond to a stimulus (a product or a brand) more strongly when it personally affects them (Zaichkowsky, 1985; Celsi & Olson, 1988). This might explain why consumers rarely express their opinions about companies' green activities. Presumably, consumers tend to overlook these activities because they are not perceived as having a direct impact on them. Indeed, Zhang, Peng, Zhang, Wang and Zhu (2014) maintained that personal life is the most popular topic in microblogging posts, accounting for 45 % of the total posts. Along these lines, Liu, Burns, and Hou (2017) found product, service, and promotions to be the dominant topics of interest to consumers when interacting with brands via Twitter.

Along with the word frequency analysis, the current study delved into the evaluation of sentiment scores for the 26 global green brands. As previously established in the literature (e.g. Lindgren, 2012; Mostafa, 2013), we found that most opinions or tweets are not affective or are ambiguous, while a minority expressed stronger sentiments. Among the brands with a prominent share of extra positive tweets were IKEA, Starbucks, and Heineken, whereas Microsoft, Heineken, and McDonald's were the brands with a higher share of extra negative tweets. In general, the number of positive tweets exceeded the number of negative tweets for all investigated green brands, with the exception of McDonald's. This underpins the findings by Jansen and others (2009) that consumers are much more likely to express positive sentiments than negative ones.

One of the indications of brands' past overall (including green) marketing efforts are also average sentiment scores, computed both in total

as well as for the positive and negative tweets. Heineken, Starbucks, ZARA, and Colgate were found to be the brands with the highest average positive score, while Nokia, L'Oreal, and Heineken exhibited the most negative sentiment scores on average. In our case, especially Heineken appeared to be the brand with significantly contrasting estimates from consumers' perspective. Looking at the overall average sentiments scores, we found Philips and Zara to be the best rated, while McDonald's was recognized as the brand with the lowest overall sentiment score. If we vaguely interpret these scores as the indicators of consumer attitudes, we might presume that, in general, the top ranked brands – Philips and Zara – are more positively perceived than McDonald's. When comparing these brands with Interbrand's (2014) perception-performance gaps, the environmental performance of Philips and Zara is perceived to be worse than it actually is, while McDonald's environmental performance is perceived to be significantly better than it actually is. It seems that a positive perception of the company's environmental activities does not necessarily result in a high sentiment score.

Additional insight into the perceptions of brands on Twitter was provided by estimating brand dispersion. For this purpose, a calculation of standard deviations and their examination in terms of their relationship to several performance measures revealed significant variations across the selected brands. Lower-dispersion (polarization or SD) brands rank higher on the Interbrand Best Global Green Brands scale, indicating a more successful performance of those brands as well as their better perception in customers' eyes than those of higher-dispersion brands. This corroborates prior findings by Luo and others (2013) about brand dispersion being negatively correlated to company performance. In addition, our study lends support to the negative relationship between high dispersion and company risk (Luo et al., 2013) by uncovering

that the stock prices of low-dispersion (or low SD) brands vary somewhat more strongly than do the stock prices of high-dispersion brands.

With respect to the polarity of tweets, several important implications emerged as a result of the present research. Practical implications for users, individuals as well as companies, include providing an overview of a general sentiment towards a specific brand or product. This could help users save time and effort of browsing through the extensive history of all posted tweets while also supporting their purchase decisions. Furthermore, sentiment analysis metrics can be used to influence brand decisions, such as brand offerings, frequency of brand messaging, timing of messaging, type of brand messaging, and brand reactions to external factors. This would allow brand managers to make better use of the Twitter service and to best influence public perception (Ghiassi et al., 2013). Although sentiment analysis cannot replace traditional measurements of customer satisfaction, such as customer surveys, it can offer additional information about customer satisfaction. In this respect, it is of particular importance to pay attention not only to mean values but also to dispersion of consumers' evaluations. In addition, companies are increasingly affected by communication in social media since customers are empowered to share product- and brand-related sentiments among each other and such exchange might strongly affect their purchase decision processes (Stieglitz & Dang-Xuan, 2013). Hence, companies might use sentiment analysis to systematically monitor user-generated content on Twitter. Along these lines, Wijnhoven and Bloemen (2014) argued that sentiment analysis can deliver important insights into the word-of-mouth (WOM) regarding products and services. This study adds to the growing body of literature on eWOM by focusing on consumer tweets. It also affords practical implications by using well-known global brands and the most widely-used microblogging site – Twitter.

6. CONCLUSION, LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

This research study contributes to the existing body of knowledge in the field of consumer microblogging behavior. The study employed a novel data analysis technique to investigate consumer sentiments toward a selection of 26 global green brands. By doing so, it elucidated several aspects of user-generated tweets that mention global green brands: their frequency, sentiment polarity and, to a limited extent, their content.

While offering an interesting springboard for future research, it also has some limitations. First, it has a limited ability to reveal consumer motivations which cannot be readily discerned through sentiment analysis. To unveil deeper drivers of consumer attitudes, intentions, and behaviors, additional qualitative insight would be required. Second, although robust in detecting basic sentiments, the lexicon-based method we used may sometimes fail to recognize subtle forms of linguistic expression (Pang & Lee, 2008). In addition, combining the lexicon-based method with other approaches, such as machine learning, might significantly improve classification performance (Dhaoui et al., 2017). Third, we have studied 26 global green brands from different sectors and, although they all fall into the category of fast-moving consumer goods, the respective companies differ in terms of their frequency of bringing new products to the market, seasonality trends, need for innovation, etc. These factors might influence the content and frequency of consumers' and companies' own tweets. Consequently, these effects should be taken into consideration in the future studies. Fourth, another limitation is the cross-sectional analysis of data collected over a longer period of time. Namely, the current study looked at the dataset collected over seven months without parsing the data longitudinally. An interesting aspect for future studies would be to connect potential deviations in sentiment scores over a period of time with companies' activities in various markets.

REFERENCES

1. Abbasi, A., Chen, H., & Salem, A. (2008). Sentiment analysis in multiple languages: Feature selection for opinion classification in web forums. *ACM Transactions on Information Systems, 26*(3), 1-34.
2. Akehurst, G., Afonso, C., & Martins Gonçalves, H. (2012). Re-examining green purchase behaviour and the green consumer profile: new evidences. *Management Decision*, *50*(5), 972-988.
3. Alaei, A. R., Becken, S., & Stantic, B. (2017). Sentiment Analysis in Tourism: Capitalizing on Big Data. *Journal of Travel Research* (upcoming).
4. Aschemann-Witzel, J., & Niebuhr Aagaard, E. M. (2014). Elaborating on the attitude-behaviour gap regarding organic products: young Danish consumers and in-store food choice. *International Journal of Consumer Studies, 38*(5), 550-558.
5. Aslam, S. (2018). Twitter by the Numbers: Stats, Demographics & Fun Facts. *Omnicore*. Available at: https://www.omnicoreagency.com/twitter-statistics/
6. Asur, S., & Huberman, B. A. (2010, August). *Predicting the future with social media*. Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology-Volume 01 (pp. 492-499). IEEE Computer Society.
7. Benhardus, J., & Kalita, J. (2013). Streaming trend detection in Twitter. *International Journal on Web Based Communities, 9*(1), 122-139.
8. Biswas, A., & Roy, M. (2015). Green products: an exploratory study on the consumer behaviour in emerging economies of the East. *Journal of Cleaner Production, 87*, 463-468.

9. Bollen, J., Pepe, A., & Mao, H. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In: N. Nicolov & J. G. Shanahan (eds.), *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, July 2011, Barcelona, Spain, (pp. 1-10).

10. Borin, N., Lindsey-Mullikin, J., & Krishnan, R. (2013). An analysis of consumer reactions to green strategies. *Journal of Product & Brand Management, 22*(2), 118-128.

11. Brand Finance (2016). *Global 500 2016: The annual report on the world's most valuable brands.* Available at: http://brandfinance.com/images/upload/brand_finance_global_500_2016_for_website.pdf

12. Bravo-Marquez, F., Mendoza, M., & Poblete, B. (2013). *Combining strengths, emotions and polarities for boosting twitter sentiment analysis.* Second International Workshop on Issues of Sentiment Discovery and Opinion Mining. ACM (Association for Computer Machinery), Chicago, IL: SenticNet.

13. Cai, K., Spangler, S., Chen, Y., & Zhang, L. (2010). Leveraging sentiment analysis for topic detection. *Web Intelligence and Agent Systems, 8*, 291-302.

14. Ceballos, M., Crespo, Á. G., & Cousté, N. L. (2016). Impact of Firm-Created Content on User-Generated Content: Using a New Social Media Monitoring Tool to Explore Twitter. In: L. Petruzzellis & R. S. Winer (eds.), *Rediscovering the Essentiality of Marketing* (pp. 303-306). Ney Jersey, NJ: Springer International Publishing.

15. Celsi, R. L., & Olson, J. C. (1988). The role of involvement in attention and comprehension processes. *Journal of Consumer Research, 15(*2), 210-224.

16. Chamlertwat, W., Bhattarakosol, P., Rungkasiri, T., & Haruechaiyasak, C., (2012). Discovering Consumer Insight from Twitter via Sentiment Analysis. *Journal of Universal Computer Science, 18(*8), 973-992.

17. Chehab, A., Liu, J., & Xiao, Y. (2016). More on intangibles: Do stockholders benefit from brand values?. *Global Finance Journal, 30*, 1-9.

18. Chen, T. B., & Chai, L. T. (2010). Attitude towards the environment and green products: Consumers' perspective. *Management science and engineering, 4*, 27-39.

19. Cordeiro, J., & Seo, J. Y. (2014). Consumer Perceptions of Green Brands and Actual Corporate Environmental Performance. In: J. Cotte & S. Wood (eds.), *NA - Advances in Consumer Research Volume 42* (pp. 443-444). Duluth, MN: Association for Consumer Research.

20. Davis, J. J., & O'Flaherty, S. (2012) Assessing the accuracy of automated Twitter sentiment coding. *Academy of Marketing Studies Journal, 16*, 35-50.

21. Decker, R., & Trusov, M. (2010). Estimating aggregate consumer preferences from online product reviews. *International Journal of Research in Marketing, 27*, 293-307.

22. Denecke, K., & Nejdi, W. (2009). How valuable is medical social media data? Content analysis of the medical web. *Information Sciences, 179*, 1870-1880.

23. Dhaoui, C., Dhaoui, C., Webster, C. M., Webster, C. M., Tan, L. P., & Tan, L. P. (2017). Social media sentiment analysis: lexicon versus machine learning. *Journal of Consumer Marketing, 34*(6), 480-488.

24. Du, R., Lu, Z., Pandit, A., Kuang, D., Crittenden, J., & Park, H. (2015, April). *Toward Social Media Opinion Mining for Sustainability Research.* Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence.

25. Eirinaki, M., Pisal, S., & Singh, J. (2012). Feature-based opinion mining and ranking. *Journal of Computer and System Sciences, 78*, 1175-1184.

26. Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data, 2*(1), 1-14.

27. Fellbaum, C. (2005). WordNet and wordnets. In: K. Brown (ed.), *Encyclopedia of Language and Linguistics.* 2nd Edition (pp. 665-670). Oxford: Elsevier.

28. Fritz, B. (2015). Tim Burton to Direct Live-Action 'Dumbo' for Disney. *Speakeasy*. Available at: http://blogs.wsj.com/speakeasy/2015/03/10/tim-burton-to-direct-live-action-dumbo-as-disney-plots-fairy-tale-strategy/

29. Ghiassi, M., Skinner, J., & Zimbra, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with applications, 40*(16), 6266-6282.

30. Ghiassi, M., Zimbra, D., & Lee, S. (2016). Targeted twitter sentiment analysis for brands using supervised feature engineering and the dynamic architecture for artificial neural networks. *Journal of Management Information Systems, 33*(4), 1034-1058.

31. González-Rodríguez, M. R., Martínez-Torres, R., & Toral, S. (2016). Post-visit and pre-visit tourist destination image through eWOM sentiment analysis and perceived helpfulness. *International Journal of Contemporary Hospitality Management, 28*(11), 2609-2627.

32. Gunter, B., Koteyko, N., & Atanasova, D. (2014). Sentiment analysis: A market-relevant and reliable measure of public feeling?. *International Journal of Market Research, 56*(2), 231-247.

33. Han, H. (2015). Travelers' pro-environmental behavior in a green lodging context: Converging value-belief-norm theory and the theory of planned behavior. *Tourism Management, 47*, 164-177.

34. Hanss, D., & Böhm, G. (2012). Sustainability seen from the perspective of consumers. *International Journal of Consumer Studies, 36*, 678-687.

35. Hartmann, P., Apaolaza Ibáñez, V., & Forcada Sainz, F. J. (2005). Green branding effects on attitude: functional versus emotional positioning strategies. *Marketing Intelligence & Planning, 23*(1), 9-29.

36. Hoepner, A. G., Dimatteo, S., Schaul, J., Yu, P. S., & Musolesi, M. (2017). Tweeting About Sustainability: Can Emotional Nowcasting Discourage Greenwashing?. *Corporate Finance, 8*(3/4), 1-18.

37. Homburg, C., Ehm, L., & Artz, M. (2015). Measuring and managing consumer sentiment in an online community environment. *Journal of Marketing Research, 52*(5), 629-641.

38. Hosta, M., & Žabkar, V. (2016). Consumer Sustainability and Responsibility: Beyond Green and Ethical Consumption. *Market-Tržište, 28*(2), 143.

39. Hsieh, C., Moghbel, C., Fang, J., & Cho, J. (2013). *Experts vs. the crowd: examining popular news prediction performance on Twitter*. In: D. Schwabe & V. Almeida (eds.), *Proceedings of 22nd International Conference on World Wide Web, ACM KDD conference. Rio de Janeiro, Brazil.*

40. Hu, G., Bhargava, P., Fuhrmann, S., Ellinger, S., & Spasojevic, N. (2017). Analyzing users' sentiment towards popular consumer industries and brands on Twitter. *arXiv preprint arXiv:1709.07434*.

41. Hu, N., Koh, N. S., & Reddy, S. K. (2014). Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales. *Decision support systems, 57*, 42-53.

42. Interbrand (2014). *The Power of Participation*. Available at: http://interbrand.com/wp-content/uploads/2015/08/Interbrand-Best-Global-Green-Brands-2014-Overview-8.pdf

43. Iyer, R. (2015). *Nike beats estimates again as demand shows no sign of abating*. Reuters. Available at: http://www.reuters.com/article/us-nike-results-idUSKBN0P52NQ20150625

44. Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American society for information science and technology*, *60*, 2169-2188.

45. Johnstone, M. L., & Tan, L. P. (2015). Exploring the gap between consumers' green rhetoric and purchasing behaviour. *Journal of Business Ethics, 132*(2), 311-328.

46. Kang, H., Yoo, S., & Han, D. (2012). Senti-lexicon and improved Naïve Bayes algorithms for sentiment analysis of restaurant reviews. *Expert Systems with Applications, 39*(5), 6000-6010.

47. Kaplan, A. M., & Haenlein, M. (2011). The early bird catches the news: Nine things you should know about micro-blogging. *Business Horizons*, *54*, 105-113.

MARKET
TRŽIŠTE

48. Kontopoulos, E., Berberidis, C., Dergiades, T., & Bassiliades, N. (2013). Ontology-based sentiment analysis of twitter posts. *Expert Systems with Applications, 40*, 4065-4074.

49. Kumar, A., & Sebastian, T. M. (2012). Sentiment analysis: A perspective on its past, present and future. *International Journal of Intelligent Systems and Applications, 4*(10), 1-14.

50. Lindgren, S. (2012). 'It took me about half an hour, but I did it!' Media circuits and affinity spaces around how-to videos on YouTube. *European Journal of Communication, 27*, 152-170.

51. Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies, 5*, 1-167.

52. Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions.* Chicago, IL: Cambridge University Press.

53. Liu, X., Burns, A. C., & Hou, Y. (2017). An Investigation of Brand-Related User-Generated Content on Twitter. *Journal of Advertising, 46*(2), 236-247.

54. Liu, Y., & Shrum, L. J. (2002). What is interactivity and is it always such a good thing? Implications of definition, person, and situation for the influence of interactivity on advertising effectiveness. *Journal of Advertising, 31*, 53-64.

55. Lu, L., Chang, H., & Chang, A. (2015). Consumer personality and green buying intention: the mediate role of consumer ethical beliefs. *Journal of Business Ethics, 127*(1), 205-219.

56. Luo, X., Raithel, S., & Wiles, M. A. (2013). The impact of brand rating dispersion on firm value. *Journal of Marketing Research, 50*(3), 399-415.

57. Miao, Q., Li, Q., & Zeng, D. (2010). Fine-grained opinion mining by integrating multiple review sources. *Journal of the American Society for Information Science and Technology, 61*, 2288-2299.

58. Minton, E. A., Kahle, L. R., & Kim, C. H. (2015). Religion and motives for sustainable behaviors: A cross-cultural comparison and contrast. *Journal of Business Research, 68*(9), 1937-1944.

59. Mocanu, D., Baronchelli, A., Perra, N., Gonçalves, B., Zhang, Q., & Vespignani, A. (2013). The twitter of babel: Mapping world languages through microblogging platforms. *PLOS ONE, 8*, e61981.

60. Montoyo, A., Martiniz-Barco, P., & Balahur, A. (2012). Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments. *Decision Support Systems, 53*, 675-679.

61. Moreo, A., Romero, M., Castro, J. L., & Zurita, J. M. (2012). Lexicon-based comments- oriented news sentiment analyzer system. *Expert Systems with Applications, 39*(10), 9166-9180.

62. Moser, A. K. (2015). Thinking green, buying green? Drivers of pro-environmental purchasing behavior. *Journal of Consumer Marketing, 32*(3), 167-175.

63. Mostafa, M. M. (2013). More than words: Social networks' text mining for consumer brand sentiments. *Expert Systems with Applications, 40*, 4241-4251.

64. Nasukawa, T., & Yi, J. (2003). *Sentiment analysis: Capturing favorability using natural language processing.* ACM 2nd International Conference on Knowledge Capture. Sanibel Island, FL, 23-25 October 2003. New York, NY: ACM.

65. O'Connor, B., Balasubramanyan, R., Routledge, B. R., & Smith, N. A. (2010). *From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series.* ICWSM Proceedings of the 4th International AAAI Conference on Weblogs and Social Media (pp. 122-129). Association for the Advancement of Artificial Intelligence, AAAI Press, Menlo Park, CA.

66. Okazaki, S., Diaz-Martin, A. M., Rozano, M., & Menendez-Benito, H. D. (2014). How to mine brand Tweets: Procedural guidelines and pretest. *International Journal of Market Research, 56*, 467-488.

67. O'Leary, D. E. (2011). Blog mining-review and extensions: "From each according to his opinion". *Decision Support Systems, 51*, 821-830.

68. Olsen, M. C., Slotegraaf, R. J., & Chandukala, S. R. (2014). Green claims and message frames: how green new products change brand attitude. *Journal of Marketing, 78*(5), 119-137.

69. Orban, P., Nagy, J., Kjarval, N., & de Carmona, X. S. (2014). Study: Exposure to brand Tweets drives consumers to take action – both on and off Twitter. *Twitter.com*. Available at: https://blog.twitter.com/marketing/en_us/a/2014/study-exposure-to-brand-tweets-drives-consumers-to-take-action-both-on-and-off-twitter.html

70. Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior, 31*, 527-541.

71. Pak, A., & Paroubek, P. (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. *Language Resources and Evaluation, 10*, 1320-1326.

72. Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval, 2*, 1-13.

73. Papista, E., Chrysochou, P., Krystallis, A., & Dimitriadis, S. (2017). Types of value and cost in consumer–green brands relationship and loyalty behaviour. *Journal of Consumer Behaviour, 17*(1), e101-e113.

74. Princeton University (2010). *What is WordNet?* Available at: https://wordnet.princeton.edu/

75. Rambocas, M., & Gama, J. (2013). M*arketing research: The role of sentiment analysis* (No. 489). Universidade do Porto, Faculdade de Economia do Porto.

76. Reyes, A., Rosso, P., & Buscaldi, D. (2012). From humor recognition to irony detection: The figurative language of social media. *Data & Knowledge Engineering, 74*, 1-12.

77. Rios, F. J. M., Martinez, T. L., & Molina, M. A. R. (2008). How green should you be: Can environmental associations enhance brand performance?. *Journal of Advertising Research, 4*, 547-563.

78. Roman, T., Bostan, I., Manolică, A., & Mitrica, I. (2015). Profile of green consumers in Romania in light of sustainability challenges and opportunities. *Sustainability, 7*(6), 6394-6411.

79. Rui, H., Liu, Y., & Whinston, A. (2013). Whose and what chatter matters? The effect of tweets on movie sales. *Decision Support Systems, 55*(4), 863-870.

80. Scheerder, J. (2015). *Running across Europe: the rise and size of one of the largest sport markets*. Boston, MA: Palgrave Macmillan.

81. Schindler, D., & Decker, R. (2013). Some remarks on the internal consistency of online consumer reviews. *Australasian Marketing Journal, 21*, 221-227.

82. Steg, L., Bolderdijk, J. W., Keizer, K., & Perlaviciute, G. (2014). An integrated framework for encouraging pro-environmental behaviour: The role of values, situational factors and goals. *Journal of Environmental Psychology, 38*, 104-115.

83. Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media - sentiment of microblogs and sharing behavior. *Journal of Management Information Systems, 29*(4), 217-248.

84. Tan, L. P., Johnstone, M. L., & Yang, L. (2016). Barriers to green consumption behaviours: The roles of consumers' green perceptions. *Australasian Marketing Journal, 24*(4), 288-299.

85. Twitter (2018). *Adding content to your Tweet*. Available at: https://help.twitter.com/en/using-twitter#adding-content-to-your-tweet

86. Wang, H. J. (2016). Green Brand Positioning in the Online Environment. *International Journal of Communication, 10*, 1405-1427.

87. Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012). *A System for Real-time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle*. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (pp. 112-120). Jeju, Republic of Korea.

88. Wang, H-J. (2017). Determinants of consumers' purchase behaviour towards green brands. *The Service Industries Journal, 37*(13/14), 105-113.

89. Wijnhoven, F., & Bloemen, O. (2014). External validity of sentiment mining reports: Can current methods identify demographic biases, event biases, and manipulation of reviews?. *Decision support systems, 59*, 262-273.

MARKET
TRŽIŠTE

90.	Wilson, T., Wiebe, J., & Hoffmann, P. (2005). *Recognizing contextual polarity in phrase-level sentiment analysis*. ACL Conference on human language technology and empirical methods in natural language processing. Vancouver, 6-8 October 2005.

91.	Wood, N. T., & Burkhalter, J. N. (2014). Tweet this, not that: A comparison between brand promotions in microblogging environments using celebrity and company-generated tweets. *Journal of Marketing Communications, 20*(1/2), 129-146.

92.	Young, M. M. (2010). *Twitter me: using micro-blogging to motivate teenagers to exercise*. International Conference on Design Science Research in Information Systems (pp. 439-448). Heidelberg: Springer.

93.	Zaichkowsky, J. L. (1985). Measuring the involvement construct. *Journal of consumer research, 12*, 341-352.

94.	Zhang, L., Peng, T. Q., Zhang, Y. P., Wang, X. H., & Zhu, J. J. (2014). Content or context: Which matters more in information processing on microblogging sites. *Computers in Human Behavior, 31*, 242-249.

MARKET
TRŽIŠTE

Appendix

Brand	Std. dev. (SD) of sentiment score[1]	Green rank 2014[2]	Perception gap in 2014[3]	Brand value rank 2014[4]	Change in brand value 2014[5]	Brand value rank 2015[4]	Change in brand value 2015[5]	Change in stock price (%) in 2014[6]	Change in stock price (%) in 2015[6]	Variation in stock price 2014[7]	Variation in stock price 2015[7]	Revenue in 2014 (in EUR mil.)[8]	Revenue in 2015 (in EUR mil.)[8]
Siemens	0.803	31	14.1	49	2%	53	-1%	-14.97	-10.9	1.37	1.28	71,227	75,636
Philips	0.842	14	5.4	42	5%	47	-8%	-19.03	-9.16	1.53	1.33	25,884	26,336
Sony	0.896	7	10.7	52	-3%	58	-5%	19.1	20.62	2.01	1.51	52,755	61,999
IBM	0.897	25	10.9	4	-8%	5	-10%	-12.2	-11.11	1.32	1.34	75,360	75,032
Canon	0.901	26	3.8	37	6%	40	-4%	6.56	-4.72	1.15	1.43	25,314	28,676
McDonald's	1.885	43	-14.5	9	1%	9	-6%	-0.05	29.75	1.18	1.37	22,286	23,327
Starbucks	1.904	37	-2	76	22%	67	16%	6.07	47.98	1.48	1.52	13,360	17,587
Zara	1.904	34	4.7	36	12%	30	16%	/	/	/	/	11,594	13,628
Pepsi	2.007	30	-2.3	24	7%	24	3%	17.06	8.59	1.16	1.4	54,156	57,881
Heineken	2.107	50	-1.5	95	-3%	89	14%	6.28	22.88	1.37	1.19	19,257	20,511

Note:

[1] Std. deviation (SD) is calculated for the sentiment scores of users' tweets which were grouped according to their brand mention.

[2] Green rank (in 2014) refers to placement on Interbrand's ranking of global green brands (Source: Interbrand).

[3] Gap in perception refers to gap between a corporation's environmental practices, and consumer's perception of those practices (negative gap = higher perception score, positive gap = higher performance score) (Source: Interbrand).

[4] Brand value rank (in 2014 or 2015) indicates the placement of brands depending on the financial performance of the branded products or services, the brand role in the purchase decision process and the brand strength (Source: Interbrand).

[5] Change in brand value (in 2014 or 2015) refers to how much brand value has changed compared to previous year, as expressed in percentage terms.

[6] Change in stock price (in 2014 or 2015) reflects how much stock price has changed within one year, as a ratio between the highest and the lowest price (Source: Morningstar Independent Investment Research).

[7] Variation in stock price (in 2014 or 2015) is calculated by comparing the highest and the lowest stock price in a given year (Source: Yahoo Finance).

[8] Revenue (in 2014 or in 2015) is obtained from the annual reports of individual companies (Revenue in 2014 (2015) is expressed in EUR, based on the exchange rate in 2014 (2015)).