Journal of Marketing Analytics (2021) 9:3–16
https://doi.org/10.1057/s41270-020-00099-z

ORIGINAL ARTICLE

Understanding changes in a brand's core positioning and customer engagement: a sentiment analysis of a brand-owned Facebook site

Zhenning Xu[1] · Colin Vail[2] · Amarpreet S. Kohli[3] · Saeed Tajdini[4]

Revised: 14 January 2020 / Accepted: 21 December 2020 / Published online: 28 January 2021

Abstract
The increasing power of social media has created unprecedented opportunities for marketers. In particular, brand-owned social media seems to be an increasingly popular way of enhancing a brand's position, connecting with customers, and improving customer engagement with the brand. To guide strategic marketing communication decision-making on social media, the current study extends the relationship communication model and offers an analytical workflow to gain new insights from unstructured textual data available on brand-owned social media. The workflow utilizes an eclectic mix of analytical tools such as word clouds, and cluster and word association analyses, which collectively allow for identification of main topics and their temporal evolution in unstructured textual data from a brand's social media. In doing so, the proposed workflow offers researchers and practitioners a step by step procedure to make sense of such textual data, which may prove unwieldy and overwhelming otherwise. Furthermore, to manifest the utility of the proposed workflow, it is applied to illustrative data collected from a brand's Facebook page. Results from this example analysis point to a slight fading of the brand's perceived core position as an event avenue, as well as an evolution of customer sentiments that may reflect different levels and types of customer engagement with the page. Finally, we discuss the implications of our findings for research and brand management practice, as well as the study's limitations and future research opportunities.

Keywords Relationship communication · NLP · Social media · Sentiment analysis · Word association analysis · Cluster analysis

Introduction

Before the era of social media, how beneficial the traditional brand communication activities are to the brand had been a key discussion point in academia for many years (e.g., Grace and O'cass 2005). Over the last few years, newer communication channels such as brand-owned social media (e.g., a brand's Facebook or Twitter pages) have become more and more popular among the brands that seek to build relationships and improve engagement with customers. Through brand-owned social media, businesses can offer support activities such as customer service, product information, special offers, and various types of entertainment (Breitsohl et al. 2015; Simon and Tossan 2018). The open nature of social media communication allows businesses to recognize the brand-consumer interactive communication as a long-term relationship investment rather than transaction-oriented interactions (Simon and Tossan 2018). Nonetheless, many studies have examined the nature of user generated content (UGC) and the many-to-many communication strategies in

✉ Zhenning Xu
zxu3@csub.edu

Colin Vail
colin.vail@td.com

Amarpreet S. Kohli
amarpreet.kohli@maine.edu

Saeed Tajdini
stajdini@ius.edu

[1] Department of Marketing & Management, California State University, Bakersfield, USA

[2] Consumer Deposits Products and Payments (CDPP), TD Bank, America's Most Convenient Bank, ME2-002-029 | 70 Gray Road, Falmouth, ME 04105, USA

[3] School of Business, University of Southern Maine, P.O. Box 9300, Portland, ME 04104-9300, USA

[4] Indiana University Southeast, Hillside Hall 41A, New Albany, USA

the online environment (Manaman et al. 2016; Wilson et al. 2012).

In the context of brand-owned social media, previous studies have offered guidance on how brands could improve customer engagement on brand-owned such media (Simon and Tossan 2018). However, there is scant research on how unstructured data from brand-owned social media could be analyzed to fine-tune a brand's communication strategy on social media and to inform insights on customer engagement. The recent boom in open source statistical platforms (R, Python, etc.) and data science tools presents an unprecedented opportunity for marketers to apply text analytical tools such as Natural Language Processing (NLP) to pinpoint main topics and topic evolutions in brand communications. For instance, recent studies applied NLP to understand service issues based on large amounts of unstructured textual data (Korfiatis et al. 2019). Nevertheless, current literature has been more focused on the volume of data or how "big" the data is (Wedel and Kannan 2016) than uncovering the embedded insights in the text (Iacobucci et al. 2019). As a result, micro-level big data has not received enough attention from researchers and practitioners (Baesens et al. 2016). Nonetheless, understanding main topics and topic evolutions on micro-level big data (e.g., brand-owned social media) has the potential to offer context-specific strategies to improve a brand's communication strategy on social media.

As a matter of fact, since the digital technologies that underpin online branding are improving rapidly, understanding how brand-owned social media might help future marketers better apply the traditional relationship communication model in a data-rich digital environment (Finne and Grönroos 2017). Marketers have long embraced the idea that "data is the new oil." However, leading tech companies such as IBM and Tableau, maintain that 'data rich but insight poor' might be the predominant dilemma faced by many marketers (Dowling 2019; Kamensky 2018), implying that big data can only be valuable if it leads to valuable information. In the current data-rich online environment, it is particularly important to understand that extracting actionable marketing insights from raw data, especially unstructured data, is challenging and painstaking. Built on the relationship communication model (Finne and Grönroos 2017), this study offers new insights on the value of analyzing brand communications on social media to improve brand positioning and customer engagement in a social media context over time and provides an analytical workflow to do so. Furthermore, different sentiment analysis methods and R packages for text analytics used in this study might open new doors for marketers to derive micro-level data-driven insights from brand-owned social media.

In sum, the main objective of this study is to propose an extended relationship communication model and an analytics workflow to aid in the analysis of brand-owned social media communication strategy. Our goal is to identify the main topics of communications on brand-owned social media and their evolutions that might be revealing of a brand's positioning on social media and the level of customer engagement, as measured by the number of likes, comments, and shares the brand's posts receive on the brand's social media. The study pursues the above objective by employing sentiment analysis methods in a workflow that transforms unstructured textual data into actionable insights. It must be noted that while customer engagement may be operationalized in numerous ways, in this study, it is measured by the number of likes, comments, and shares that a brand's posts receive on the brand's social media (Lee et al. 2018). Data collected from the Facebook page of a popular event marketer located in the Northeastern United States was used to provide an exemplar application of the proposed workflow. Insights from the analysis led to discovery of main topics, situational factors, topic evolutions, and a decrease of customer engagement (Finne and Grönroos 2017; Simon and Tossan 2018). Figure 1 is a graphical representation of the phenomenon that the current study investigates. This figure features a feedback mechanism linking analytics, communications, and brand positioning and customer engagement on social media.

The paper begins with an introduction of the topic, followed by a review of the relevant literature including brand positioning and brand communication in the social media context, the relationship communication model, and data mining of social media. This is succeeded by the methodology and the analysis sections. Next, a discussion of the results details how the relationship communication model was examined based on brand-owned social media text. A

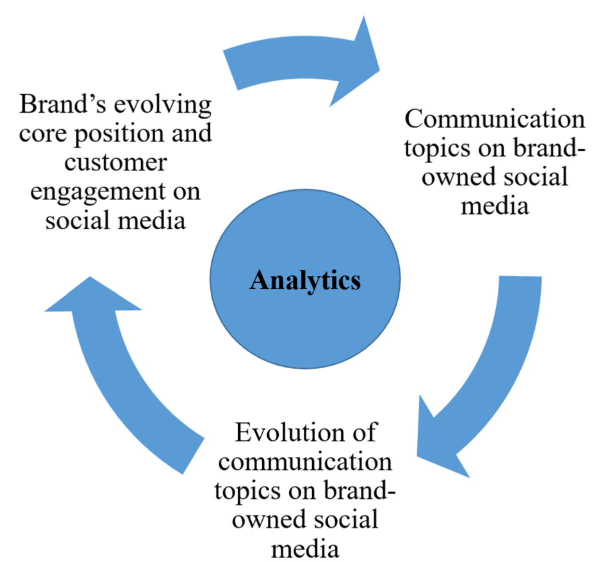


Fig. 1 A feedback mechanism linking analytics, communications, and brand positioning and customer engagement on social media

section detailing our conclusions follows, and the paper ends with a discussion of the research limitations and future research opportunities.

Literature review

Sentiment analysis of marketing data

Traditionally viewed as a tool exclusive to marketing communications purposes, social media is a potentially useful for marketing management in general (Bukhari et al. 2015). Of course, the use of social media analytics has not been limited to the field of marketing. For example, in crisis and disaster management, social media content has been mined and utilized to investigate real-time phenomena (Pohl et al. 2016; Wang and Ye 2018).

Shao (2009) argues that online activities include consumption of information, entertainment, engaging in social interactions, developing online communities, and self-expression through content creation. The latter two provide major opportunities to evaluate customer sentiments. Sentiment analysis is crucial in managing brand communications since sentiments (both user generated content and company generated content) have been shown to influence customer-brand relationships (e.g., Lopez et al. 2020). As such, given the volume of online social media data, sentiment analysis has become increasingly popular in marketing research (see Rambocas and Pacheco 2018 for a thorough review of the application of sentiment analysis in marketing research). Liu (2012) defines sentiments as subjective feelings and thoughts, and Rambocas and Gama (2013) define sentiment analysis as the application of machine learning techniques to evaluate and classify attitudes and opinions on topics of interest.

Perhaps the application of sentiment analysis has never been more accessible than now as social media such as Facebook have introduced new customer-centric tools that enable customers to follow and interact with their favorite brands (Baccarella et al. 2018), form communities around, and express their attitudes toward those brands, leaving a trail of highly valuable data online. It is well recognized that consumers' following of a brand's social media account indicates loyalty and emotional attachment to the brand (Simon and Tossan 2018). With the rapid growth of online data, companies now have opportunities to gather data from diverse stakeholders using various tools. In addition, there is now an increasing push to connect with consumers not only through one-to-one online communication strategies, e.g., online surveys, but also through many-to-many communications, e.g., chat rooms, blogs, and discussion forums (Maclaran and Catterall 2002).

Brand positioning and customer engagement on social media

"A brand is a combination of features (what the product is), customer benefits (what needs and wants the product meets) and values (what the customer associates with the product)" (Pearson 1996, p. 6). The concept of brands and branding is of highest importance in marketing. This is because, as Keller and Lehmann (2006) note, brands are directly related to a company's customer, product, and financial bottom lines.

As a related concept, brand positioning refers to marketing activities that establish specific associations about the brand in the customer mind that distinguish the brand from competitors and give it competitive superiority (Keller 2002). Keller and Lehmann (2006) consider three major ways that, in conjunction, can be used to achieve the desired brand positioning: (1) integrating brand elements, (2) integrating marketing channels and communications, and (3) combining company-controlled and external events. Brand elements include, but are not limited to, the brand name, packaging, logo, and symbols. Integrating brand elements means choosing and designing these elements in alignment with a unified image of the brand. As for integrating marketing channels and communications, what it meant here is orchestrating different means of communication (e.g., TV and print ads, trade and consumer promotions, public relations, direct response ads, sponsorships, etc.) and various channels (e.g., retailers, online stores, company-owned stores, catalogues, telephone or mail orders, etc.) to create a synergistic confluence on the brand's projected position. Lastly, by combining the customer experience from company-controlled events (e.g., sweepstakes) and external venues (e.g., independent online brand forums), companies can leverage the advantages of each type of event to fortify the desired brand experience. However, whether the brand communication activities work in a brand's favor is still unknown. The current study strives to expand our understanding of the second and third brand positioning strategies mentioned above.

It is commonly accepted that a brand positioning strategy is no longer created solely by companies or marketers; rather, it usually involves creating bidirectional communications and interactions between brands and consumers (Finne and Grönroos 2017; Rosenthal and Brito 2017). The nature of such communications on social media means that businesses have less controllable brand meanings (Rosenthal and Brito 2017). Surprisingly, the focus of social media has been on promotion and conveyance of brand meanings per se for a long time. The "data rich but insight poor" dilemma suggests that for many years, little emphasis has been placed on how to generate actionable, data-informed insights from social media platforms using social media analytics (Dowling 2019; Kamensky

2018). Much has changed since the rise of social media and smartphones during the 2010s. These changes have occurred in three broad areas: (1) the audience (social media users as opposed to the general public), (2) the development tools and APIs working as interfaces between the social media and data, (3) more intuitive and practical analytics platforms and tools available to marketers and data analysts (Google Analytics, Tableau, Salesforce, R, Python, and dedicated R and Python libraries, etc.).

The evolution of social media and the plethora of data attached to it have numerous marketing applications. Historically, there has been a focus on how companies could implement, monitor, and measure customer relationships (Lindgreen and Crawford 1999). Kelly and Fairley (2018) suggest that it is critical for marketers to continuously strive for better brand positioning strategies and enhanced relationships with customers. Doing so has become more complicated in our current environment where diverse information flows in from different social media as well as traditional marketing channels. Engaging customers with a brand's social media content is an especially critical goal of relationship management since such customer engagement is conducive to critical marketing outcomes such as word-of-mouth intentions, among other aspects of consumer buying behavior (Chan et al. 2014).

Building effective brand positioning and customer engagement strategies can no longer be approached from a traditional perspective and must rather be tailored to the digital environment, the data-rich nature, and a company's specific industry and competencies (Baesens et al. 2016). Furthermore, companies are expected to establish feedback mechanisms to ensure that their strategies are subject to actionable insights based on data from different communication channels.

Brand-owned social media and its interface with brand positioning and customer engagement

In the current social media era in which both marketers and consumers are engaged in the process of cocreating brand identity, the development of new ways of using big data analytics becomes more critical than ever (Finne and Grönroos 2017; Xu et al. 2015). Unlike UGC, content from brand-owned social media is controlled by corporate channels (Camiciottoli et al. 2014; Edelman 2010). To gain deeper managerial insights, several pioneer studies recommend that brands take a proactive stance on analyzing consumer reactions to UGC and content from brand-owned social media. Since not all content from brand-owned social media are created equal, an investigation of the value of such content on branding is critical (Gensler et al. 2013).

Communication-in-use, value-in-use, and the relationship communication model

The notion of communication-in-use postulates that various information-gathering and decision-making activities centered around communications are designed to create meaning reflected as experience, processes and activities and contribute to our understanding of the marketing ecosystem (Finne and Grönroos 2017). According to the communication-in-use perspective (Finne and Grönroos 2017), the final value of marketing communications will depend on our ability to fine tune our marketing messages ("communication-in-use") and to make sense of them ("value-in-use").

According to the relationship communication model, the consumer's perception of brand communications is complex and depends on several situational factors. These situational factors must be taken into consideration in firms' communications to ensure that the firm and its audience are on the same proverbial page. For example, if a firm's message contains the notion of long-term goals, the audience may have a very different perception of what constitutes "long term" than what the firm intended. Further complicating the phenomenon, these situational factors may vary among the individuals in the audience as well, requiring a segmentation approach to communications. Situational factors can be classified into historical factors, future factors, internal factors, and external factors (Finne and Grönroos 2009, 2017). While historical and future factors relate to the consumer's perception of a given relationship, external and internal factors relate to the consumer's individual context (Finne and Grönroos 2009). Historical factors include messages, memories, and stories from the past that impact the consumer's meaning creation with regards to the brand. Future factors relate to the expected future of the consumer-brand relationship (Edvardsson and Strandvik 2000). External factors include the economic situation, culture, familial, and competitors' messages in the surrounding society.

Finally, internal factors include identity, capabilities, and abilities, and personal interests. In the context of consumers' social media interactions with brands, these factors are expected to surface in the form of frequently used words and expressions, which translate into measurable customer engagement metrics on social media due to the feedback mechanism as defined by the relationship communication model (Finne and Grönroos 2017) These factors will be further discussed in the analysis section of the paper.

De Chernatony (1999) makes the argument that brand identity should be created internally by taking measures to align employee actions with the company's positioning strategy. While this is especially true when brands only rely on traditional media, marketers emphasize the influence of historical and future factors and the way how social media creates a feedback mechanism when communication messages are delivered through social media (Finne and Grönroos 2017). Since consumer attitudes and behaviors (e.g., likes, comments, and shares in the context of social media) are constantly evolving, taking a process perspective instead of an outcome view is more important than ever (Finne and Grönroos 2017). This is why the current study focuses not only on identification of topics on social media posts but also their evolution over time (see Fig. 1).

To deepen our understanding of the connections between data, information, and knowledge, Krishen and Petrescu (2018) recently introduced a framework that emphasizes the critical role of analytics in new directions for research development. This framework recommends an analytics-oriented approach to understand the process of converting data to information. Although the framework is for research development in marketing analytics, it offers a creative way for understanding the way analytics works in the new marketing analytics paradigm. This study continues the same theme by proposing an adapted relationship communication framework (see Fig. 2) for integrating the analytics view into the original relationship communication model proposed by Finne and Grönroos (2009, 2017).

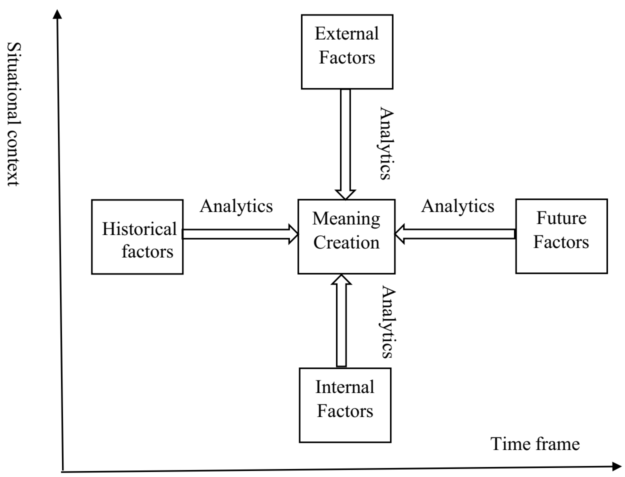


Fig. 2 An integrated relationship communication model that connects situational factors to meaning creation via analytics (adapted from Finne and Grönroos 2009, 2017)

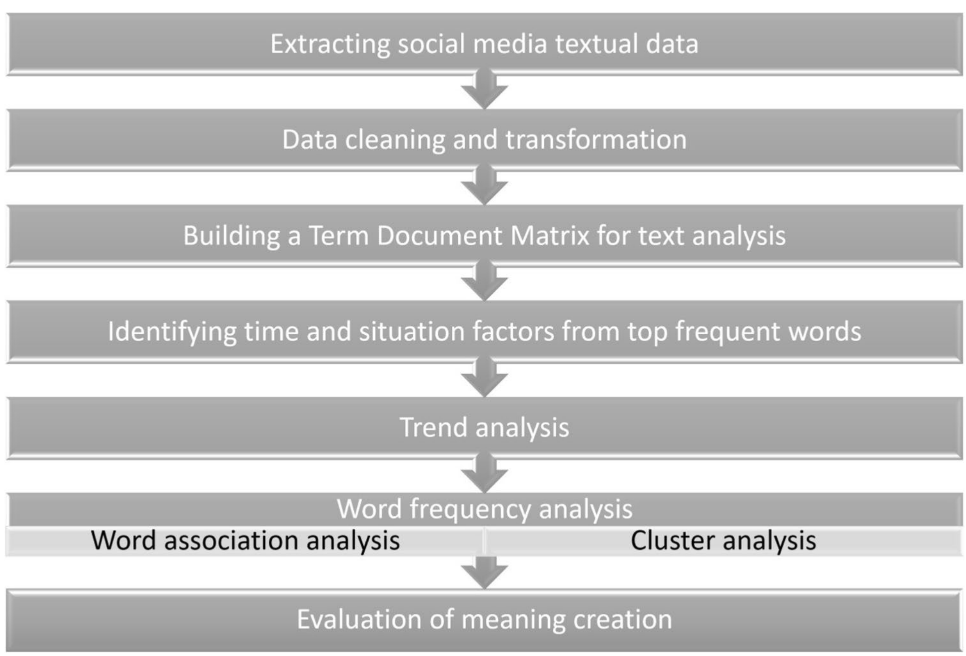
To recap the discussion so far, finding new ways of understanding the value and the evolution of topics in brand-owned social media content can open a novel door to the understanding of brand positioning and customer engagement on social media. Fortunately, the "open source" nature of the digital environment and the recent surge of new text-analytics techniques make it easier for brands of all sizes and resources to achieve this goal. Against this backdrop, the current study is an attempt to expand the current literature by modifying the original relationship communication model proposed by Finne and Grönroos (2009, 2017).

Additionally, the study offers an implementable workflow (Fig. 3) that employs a mix of analytical tools to better understand and manage two of the most critical marketing outcomes—brand positioning and customer engagement.

Data

We collected all posts on a multi-purpose event avenue's own Facebook page from 2015 to 2016. The chosen company is a multi-purpose event venue located in a small New

Fig. 3 Text analytics workflow for evaluating brand-owned social media textual data



England port city. The property is home to an outdoor concert venue and other local vendors. The Facebook posts were retrieved in JSON format using Facebook's developer API (https://developers.facebook.com) and was subsequently manipulated into spreadsheet format using the R language and several dependent NLP packages (e.g., Rfacebook). While media content, such as videos and images, could not be captured, a subset of related data (likes, reactions, etc.) was recorded for all the posts. The observations contained both text (e.g., user comments) and numerical values (e.g., a post's number of likes).

Using the Facebook API, we extracted 408 Facebook posts from the company's Facebook page, which at the time of the data collection, constituted all of the posts by the company. These Facebook posts represent brand-owned social media content, as opposed to user generated content in most studies. Brand-owned social media content is controlled by corporate channels and is made available on social media platforms (Edelman 2010). The information gathered on user interactions (shares, likes, and comments) with each post constitutes a rudimentary evaluation system by which marketing effectiveness can be gauged. As mentioned previously, this is how the current study operationalizes the notion of customer engagement with the company's social media content.

To address the methodological gaps for evaluating brand-owned social media content using text analysis and NLP, we integrate several methods and develop a novel workflow to help future researchers and practitioners explore the benefits of text analytics for managing brand communication and customer engagement (see Fig. 3).

Methodology

Trend analysis

As it can be seen in Fig. 3, the first analysis in the proposed workflow is trend analysis. To understand how the brand's Facebook posts affect the relationship building in years 2015 and 2016, we first examine the relationship between the Facebook posts and customer engagement, as measured by the posts' average number of likes, comments, and shares. Figure 4 shows different levels and types of customer engagement in 2015 and 2016, and how customer engagement can be disproportionate to the brand's content productivity, as measured by the number of posts. Counterintuitively, we found that while the brand's content productivity had increased by 70.3% from 2015 to 2016, there was a significant decrease in engagement as measured by the average number of likes (17.71 vs. 10.54), the average number of comments (4.5 vs. 3.15), and the average number of shares (1.45 vs. 0.42), in the same time period.

While the why of it (decreased customer engagement) is outside the scope of this study, this sizeable decrease in customer engagement warrants an investigative action by the brand.

Specifically, given the empirical association between the decreased customer engagement and the shift in the main topics, their evolutions, and the brand associations, a researcher may hypothesize causal relationships and further investigate its direction and other attributes.

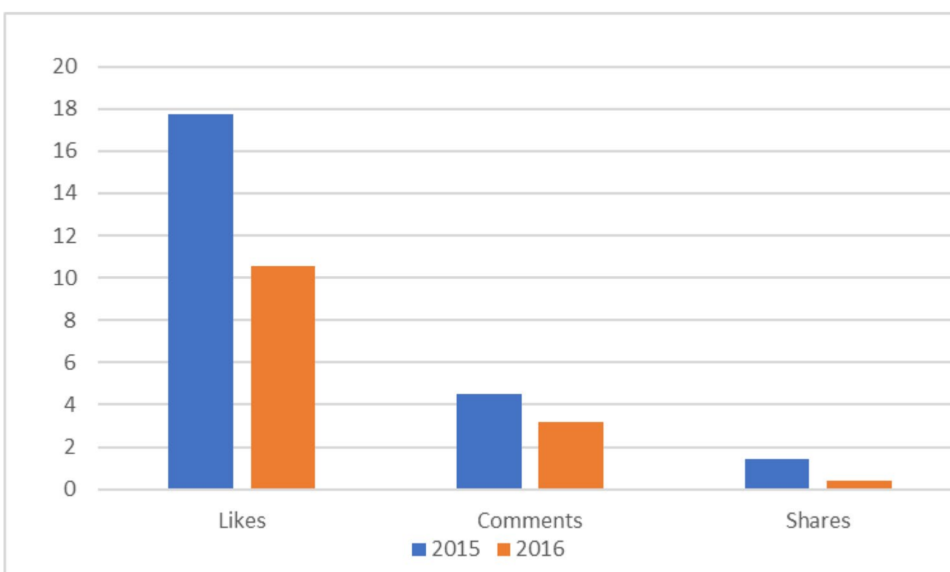


Fig. 4 The average number of likes, comments, and shares in 2015 and 2016

Sentiment analysis

Of importance to our research is to identify topic evolutions in brand-owned social media textual data. Using R statistical software and packages, we identified Facebook posts as being originated in 2015 or 2016, respectively, and coded them as two separate word corpora. The text was then cleaned of punctuation, special characters, and blank space and then transformed into lower case. From the corpus generated, we created a matrix listing each unigram (word by word text) along with the total amount of usages by year.

Word clouds are used to visually summarize text documents. Figure 5 was generated based on the word corpora using the WordCloud package available in R. Word size reflects the frequency of different words by year. For instance, larger words towards the center of the cloud reveal the most frequent words of the brand-owned social media data published in 2015 and 2016. While the interpretation of word clouds, as a qualitative tool, may be subjective, it must be noted that this tool is used in the proposed workflow in conjunction with other more quantitative and objective tools such as cluster analysis, creating a more comprehensive and multi-dimensional portrait of the topics and their evolution over time.

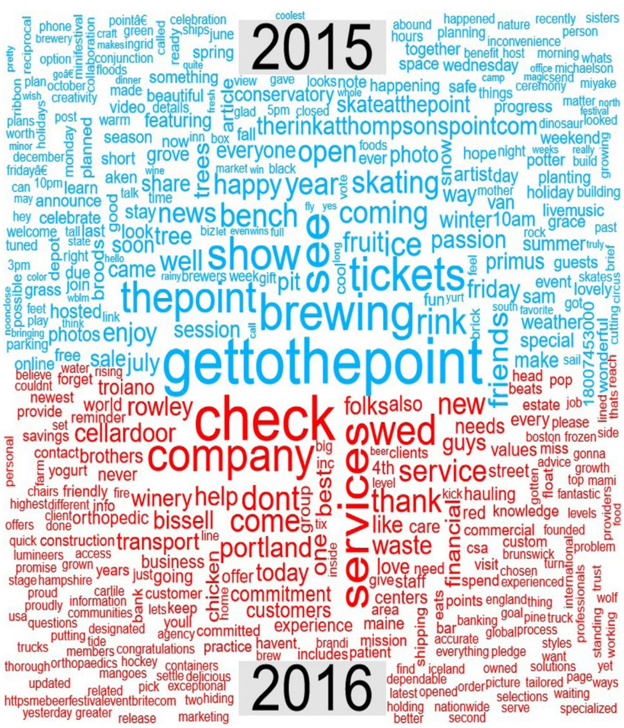


Fig. 5 Comparison of word clouds featuring Facebook post content in 2015 and 2016

Figure 5 shows that the company's only attempted hashtag—gettothepoint—is highly prevalent in the 2015 corpus. Additionally, the word clouds show frequently used words such as 'theater,' 'thepoint,' 'brewer,' and 'rink,' which seem to be aligned with a projected brand position entailing recreational activities such as drinking and watching movies.

To further understand the reason of decreased customer engagement and the topic evolutions in the word clouds across years 2015 and 2016 presented in Fig. 5, we identified the top three situational factors from the top ten frequent terms from each corpus (see Table 1) and listed the top ten associated topic words that appear with the top three situational factors in each year (see Table 2). According to the original relationship communication model (Finne and Grönroos 2017), situational factors are the aspects that often influence the consumer's meaning creation process. The effectiveness of these situational factors in brand communication could manifest itself in the way how customers interact with the virtual touchpoints, as reflected by the number of likes, comments, and shares of the social media posts in the context of this study. For instance, since the place is named "Thompson's Point," the most important situational factor for both years is "point."

Table 2 illustrates which top 10 topic words are used most often in combination with the top three situational factors in 2015 (i.e., "point," "brewing," and "tickets") and 2016 (i.e., "point," "services," and "theatre"), respectively. Note that multiple minimum support rules for word association analysis were developed. These minimum thresholds are to be determined by the individual researcher according to the context of the study (Pedersen and Kolhatkar 2009). The minimum threshold used is 0.3. For comparison purposes, if a threshold of 0.3 were used, then there would be more than

Table 1 Top situational factors identified using the word frequency analysis

2015	Point	Brewing	Tickets	Show	Rink	Summer	Project	Theatre	Bench	Friends
Frequency	34	23	17	15	15	14	11	11	10	10
2016	Point	Services	Theatre	Bissell	Brothers	Folks	Summer	Project	Beer	Brewing
Frequency	38	18	13	12	12	11	11	11	11	9

This table features top ten situational factors identified in 2015 and 2016, respectively. The top three situational factors for 2015 (i.e., "point," "brewing," and "tickets") and 2016 (i.e., "point," "services," and "theatre") were used for the word association analysis in Table 2. *Note* common words (or 'stop words' like the, are, is, day, life, etc.) were removed to eliminate noise or distracting features

Table 2 Word association analysis—top topic words associated with top situational factors

Year 2015						Year 2016					
Top words associated with "point"	Association	Top words associated with "brewing "	Association	Top words associated with "tickets "	Association	Top words associated with "point"	Association	Top words associated with "services"	Association	Top words associated with "theatre"	Association
Thompsons	0.59	Allagash	1	Sale	0.58	Thompsons	0.64	Including	0.64	State	1
Rink	0.41	Andrews	1	Show	0.56	Cellardoor	0.41	Troiano	0.62	Head	0.47
Theatre	0.38	Atlantic	1	Online	0.5	Winery	0.38	Convention	0.55	Page	0.47
State	0.37	Banded	1	Conjunction	0.46	Destination	0.32	cvb	0.55	Music	0.47
Show	0.34	bAxter	1	Box	0.46	Monthly	0.32	Information	0.55	Please	0.42
gettothepoint	0.33	bay	1	office	0.46	performances	0.32	Marketing	0.55	Contact	0.41
New	0.33	Bear	1	Person	0.46	Opening	0.3	Motorcoach	0.55	Concert	0.41
Conservatory	0.33	Bigelow	1	Phone	0.46			Relations	0.55	Tickets	0.41
wwwthomp-sonspoint-mainecom	0.33	Bottling	1	Options	0.46			Tourism	0.55	Info	0.41
Take	0.33	Bow	1	Advance	0.44			Tourists	0.55	Questions	0.39

Notes Table 2 illustrates which top ten topic words are used most often in combination with the top three situational factors in year 2015 (i.e., "point," "brewing," and "tickets") and 2016 (i.e., "point," "services," and “theatre”), respectively. Situational factors are the aspects that often influence the consumer’s meaning creation process and the way how customers interact with the virtual touchpoint (e.g., likes, comments, and shares in the context of this study). This table showcases a word association analysis for three of the top frequent words in 2015 (i.e., "point," "brewing," and "tickets") and 2016 (i.e., "point," "check," and “theatre”). The lower panel of the table illustrates which topic words are used most often in combination with the top three situational factors for each year, respectively. Note that due to space limitation, only the top ten associated words are listed in the lower part of the table and the words are kept in their verbatim form, and thus may contain typos, abbreviations or acronyms

10 topic words that are associated with the most important situational factor that represent the core brand message (i.e., "point") in 2015, while there are only 7 topic words that are associated with "point" in 2016. Due to space limitations, only the top 10 associated topic words were included. However, in 2016, there are less than 10 associated topic words for the word "point." In general, the analysis shows that there are more topic words being used in conjunction with the top three situational factors in year 2015 than there are in year 2016.

The word clouds and word frequency analysis for 2016 show common words that project the brand's generic role as a service provider (service(s), theater, summer, project, etc.), which contrast starkly with the common words of 2015 that project the brand's more nuanced role as an event venue (brewing, tickets, show, etc.). The major topics behind the brand-owned social media textual data in 2015 could be considered critical to the central branding strategy. In the discussion section, we briefly discuss the application of the relationship communication model to our analysis and offer managerial insights.

Hierarchical cluster analysis of Facebook sentiments

To better compare changes in brand positioning from 2015 to 2016, we take advantage of hierarchical clustering. The purpose of performing this analysis is to classify the main topics and to showcase the related topic patterns in each year. Building on what has been extracted from analyzing term frequencies and correlation, clustering categorizes text documents and classifies the representative terms for each text document (Maceli 2016). Using the corpora comprised of the words from Facebook posts, we create a term document matrix (TDM) that lists unigrams within the corpus as rows.

Hierarchical clustering was performed using the Complete Linkage Method with recomputed distances calculated based on the Lance Williams dissimilarity metrics. The clustering unit is words found in the Facebook posts. The number of individual clusters was set to four. To create the individual clusters, the model identifies the 4th level of dissimilarity from the final merging of similar clusters, visualized as the top of the dendrograms. As a note for interpretation of the dendrogram, the algorithm used assigns tighter clusters to the left of each subtree within the dendrogram.

The resulting clusters identify how the topics are grouped. The clusters are numbered from left to right in Figs. 6 and 7. Clusters 2 and 3 can be said to capture the brand's meaning creation process as a form of situational factors as these clusters contain words related to the company's main attributes such as its location and the product offerings. Cluster 4 contains what we believe supplements clusters 2 and 3 as the various services and events that support the brand's meaning creation process.

As specified in the workflow, the focus of this step is to make sense of the relationship communication process (Finne and Grönroos 2009) using sentiment analysis. The 2015 cluster dendrogram shows the brand's situational factors are captured in clusters 1, 2, and 3 as these clusters contain words related to the company's service offerings, location information and

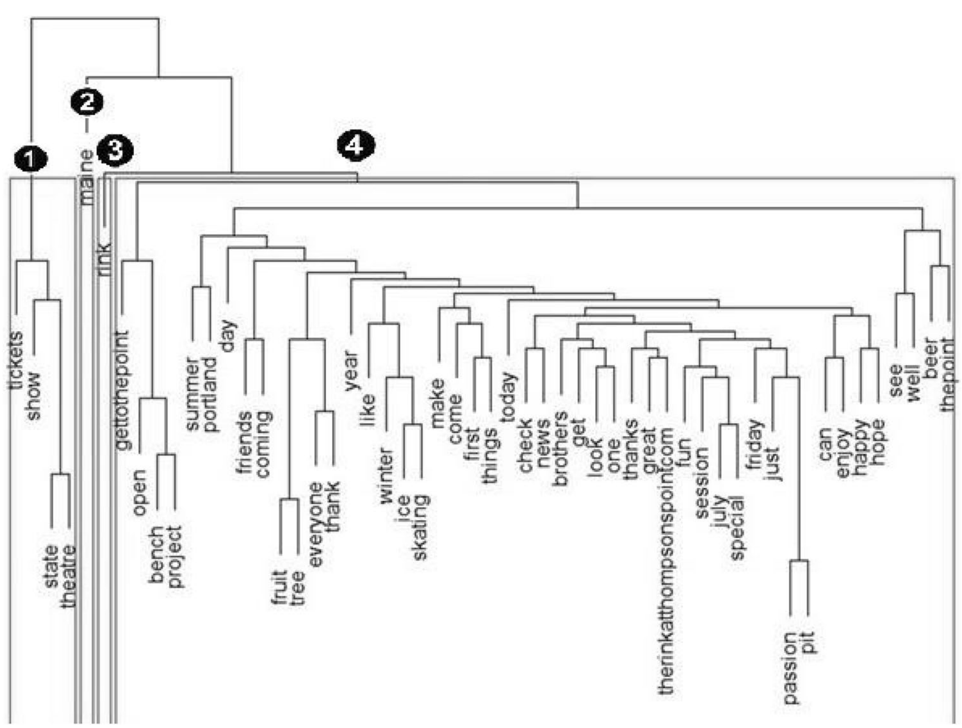


Fig. 6 Hierarchical clustering dendrogram ($k = 4$) constructed from the 2015 Facebook post content. *Note* the numbers at the top of the main dendrogram branches denote cluster numbers. In addition, the clustering subtree shows how words found in the Facebook posts are being grouped together

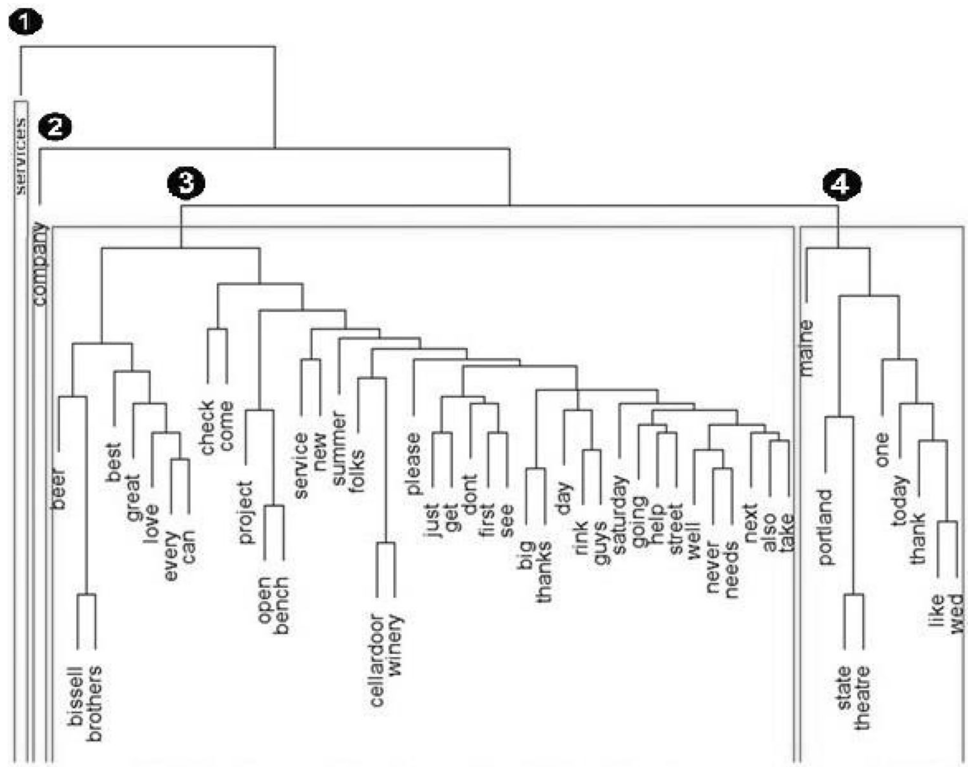


Fig. 7 Hierarchical clustering dendrogram ($k = 4$) constructed from the 2016 Facebook post content. *Note* the numbers at the top of the main dendrogram branches denote cluster numbers. In addition, the clustering subtree shows how words found in the Facebook posts are being grouped together

brand-related features. For example, "tickets" (1) and "rink" (3) clearly convey what services are offered and are consistently used in messaging based on their locations within the dissimilarity matrix. Location information of the venue is an important part of owned content. The term "Maine" (2) is the highest associated term. Cluster 4 contains primarily descriptive terms that make up the tone and attitude within owned media; there is also evidence of specific branded terms used frequently enough to be part of the analysis.

In contrast, clusters from 2016 reveal a different set of situational factors in different clusters that are not closely aligned with the brand's primary brand position. In addition, the two tightest clusters (least amount of dissimilarity) located on the left side of the dendrogram are composed of two terms only ("services" and "company"). These words are associated with a broader context of the business vs. the venue's specific offerings. The relationship building cluster (3) contains many similar terms to the same section in the previous year, although the brand specific terms are no longer frequent enough to be represented in the clustering analysis.

Discussions

While the number of total Facebook posts in 2016 is 37% greater than the previous year (236 vs. 172) in our data, the content of those posts contains very few terms that are associated with the brand's core positioning and relationship building. According to our word association and cluster analyses, the main relationship building strategy is to create strong brand associations with time and situational factors (Finne and Grönroos 2009). In 2015, there are recognizable core brand messaging centered around the venue's brand such as "gettothepoint" and "thepoint" that are derived directly from the location name. The social media presence in 2015 contains elements of a burgeoning brand strategy driven by name recognition and location specific services (tickets and brewing) that appears to have diminished by the following year.

When the communication strategies are compared, we see relative similarities but less depth in 2016. Our comparative word cloud and word association analyses show that there was a slight fading of the brand's core brand position and that the communication language used in 2016 is focused on a broader context of business activities, thus less engaging and interactive with consumers. Our cluster analysis reinforces the patterns shown through the comparative word cloud and word association analyses while adding a way to group and identify how different terms are linked to the company's core branding messages. It also documents how in a relatively short period of time the company's social media content shifted in a significant way based on the frequency and combination of the terms within the messaging. These visual representations of the company's brand-owned social media data may act as an internal brand position mechanism or feedback loop through which the company can employ and quantify relationship communication strategies that target the specific channels and interests of the identified segments.

Conclusions and contributions

User generated content (UGC) and the many-to-many communication strategies in the online environment have been widely studied (e.g., Manaman et al. 2016; Wilson et al. 2012). Recently, brand-owned social media have risen in popularity among brands that strive to improve their perceived brand positioning and customer engagement. While there has been significant academic strides to better understand and utilize the data produced on such media, there is scant research on how main topics and their evolutions could be identified from brand-owned social media to inform strategic insights on brand positioning and customer engagement. However, the advent of affordable and applied data science tools presents a golden opportunity for companies to achieve this goal. But as mentioned previously, the current literature has been more occupied with the volume of data or how "big" the data is (Wedel and Kannan 2016) than discovering the deeper meanings in textual and micro-level big data such as brand-owned social media data (Baesens et al. 2016; Iacobucci et al. 2019). Notwithstanding this, understanding the topics and their evolutions in such data has the potential to offer context-specific insights to enhance brand positioning and customer engagement. The current study is an attempt to fill this gap by offering marketing practitioners and researchers with an intuitive, analytical workflow to achieve this goal. Moreover, to exemplify the utility of the proposed workflow, we apply it to data from a brand's social media to make conclusions about the brand's perceived positioning and customer engagement.

To better understand the brand–customer communications on brand-owned social media, we expanded the Relationship Communication model (Finne and Grönroos 2009). This model considers four situational factors influencing the process of meaning creation (Finne and Grönroos 2009). According to this perspective, the meaning creation process goes beyond the traditional way of communication by considering the importance of time and situational factors. Considering that there is scant empirical research regarding the impact of relationship communication using brand-owned social media textual data, this study offers a timely contribution to advance our understanding of the relationship communication model in the current "data rich but insight poor" marketing environment.

Traditionally, it has been difficult to evaluate how brand-owned social media sites engage with the customers. Our analyses suggest that the decrease in engagement and pivots in corporate content should not be viewed as separate events. According to the traditional IMC model, the corporate content is usually "pushed" without actionable data for a recursive communication feedback loop. The relationship communication model as proposed by (Finne and Grönroos 2009) recommends a recursive loop that emphasizes the dynamic relationship in which customers co-create value in the joint sphere (e.g., a social media channel such as Facebook) where the brand's and the customer's value creation activities merged into a single process (Lenka et al. 2017).

The comparative word clouds, word frequency analysis, word association, and cluster analyses are incorporated in the proposed analytical workflow to visually illustrate how brand-owned content, in the form of social media textual data, can be used to evaluate brand positioning and customer engagement over time. Insights from this form of analysis become more effective when a relationship communication strategy exists, and where core brand messaging (i.e., hashtag 'gettothepoint') and other situational factors (e.g., 'theater') can be communicated and tracked over time.

To sum, when it comes to long-term strategic communication planning and social media marketing management, brands usually place more emphasis on UGC. However, brand-owned social media communication offers companies more control and opportunities for long-term brand positioning and customer engagement. Given the current shift from traditional marketing to relationship marketing, this study proposes an integrated relationship communication model and an intuitive workflow that leverages the power of text analytics to learn more about the brand-customer relationship. In the example Facebook data analyzed, the analysis of topics revealed a significant change in customer engagement over time which may be related to the identified change in the brand's perceived positioning over the same time period. Such findings may sound the alarm for the brand management to conduct follow-up studies on potential causal analysis and on how to remedy the situation. Moreover, this emphasizes the importance of having a longitudinal view of social media marketing and incorporating in it a feedback mechanism.

Future research and limitations

The main offering of this paper is the extended relationship marketing model in Fig. 1 and the analytical workflow proposed in Fig. 3. The relationship communication model offers useful insights and a theoretical framework on strategic planning in communication and branding.

However, to the best of our knowledge, the model has not been tested with empirical data. This study suggests that the model is beneficial for managers who are looking forward to adopting text-analytics techniques in an actionable, step by step framework.

Using extended relationship marketing model and the workflow, marketing managers can make sense of unstructured data from their social media, specifically to better understand their market-perceived core brand position, as well as the levels and types of customer engagement they are garnering on social media, based on the identified topics in the data and their evolutions.

The case study related to the event marketer acts as an example to demonstrate how the proposed workflow can be applied. While we do not have any reason to believe that the framework's applicability should be limited to the demonstrated example, we acknowledge it as a limitation of the current study that the case of only one brand was examined. Different brands may exhibit different profiles when it comes to the number, type, depth, and speed of evolution of topics on their social media, and these factors may impact the utility of our proposed analytical workflow for different brands.

Another limitation of the current study relates to the data used as an example. Having used Facebook data only may have limited the generalizability of our results. Brands may own different types of social media and communications on each medium may exhibit different nature. Also, different social media provide different data that can be used as variables by marketers. For example, Facebook may offer a more versatile set of metadata on content that can be used to measure customer engagement than YouTube. A follow-up study that applies the proposed framework to data from different social media may increase the generalizability of our results. In addition, the extraction and sense-making of such unstructured data may sound challenging and complex. Moreover, marketers may not see immediate benefits for extracting and analyzing the brand-owned content from their social media. More structured forms of market data (surveys, focus groups, lab observations, etc.) may sound more appealing and straightforward to marketers, especially those without the required infrastructure to collect and analyze large amounts of unstructured data.

In our view, one strength of the proposed framework is that it integrates a number of more qualitative and more quantitative tools into a process that takes in raw data and outputs insights. On the flip side, the qualitative and subjective nature of some of the analyses and interpretations may feel uncomfortable to more objective researchers and practitioners. As specific examples, the comparisons of the themes of the word clouds over time, as well as, identifying the themes of the clusters from the dendrogram both rely heavily on the researcher's intuition, domain and customer knowledge.

Finally, one of the major challenges facing many marketers today is how to integrate different information from different channels for strategic decision-making. Many organizations would side with a data-driven approach. Previous literature suggests that theory and empirical analysis go hand-in-hand (Ågerfalk 2014). Our workflow and a micro-level big data analysis can assist in achieving this goal and offer a complementary approach to show that a data-driven approach is not enough. Future research could investigate the given size and resources a company would need to successfully implement the use of modern big data analytics within their communication and branding strategy.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Ågerfalk, P.J. 2014. Insufficient theoretical contribution: a conclusive rationale for rejection? *European Journal of Information Systems* 23 (6): 593–599.

Baccarella, C.V., T.F. Wagner, J.H. Kietzmann, and I.P. McCarthy. 2018. Social media? It's serious! Understanding the dark side of social media. *European Management Journal* 36 (4): 431–438.

Baesens, B., R. Bapna, J.R. Marsden, J. Vanthienen, and J.L. Zhao. 2016. Transformational issues ofbigdataand analytics in networked business. *MIS Quarterly* 40 (4): 807–818.

Breitsohl, J., W.H. Kunz, and D. Dowell. 2015. Does the host match the content? A taxonomical update on online consumption communities. *Journal of Marketing Management* 31 (9–10): 1040–1064.

Bukhari, I., C. Wojtalewicz, M. Vorvoreanu, and J.E. Dietz. 2015. Social media use for large event management: The application of social media analytic tools for the Super Bowl, In *2012 IEEE Conference on Technologies for Homeland Security (HST)* (pp. 24–29). IEEE.

Camiciottoli, B.C., S. Ranfagni, and S. Guercini. 2014. Exploring brand associations: An innovative methodological approach. *European Journal of Marketing* 48 (5–6): 1092–1111.

Chan, T.K., X. Zheng, C.M. Cheung, M.K. Lee, and Z.W. Lee. 2014. Antecedents and consequences of customer engagement in online brand communities. *Journal of Marketing Analytics* 2 (2): 81–97.

De Chernatony, L. 1999. Brand management through narrowing the gap between brand identity and brand reputation. *Journal of Marketing Management* 15 (1–3): 157–179.

Dowling, Z. 2019. Data rich but insight poor: How companies are missing the point. *Customerthink*. https://customerthink.com/data-rich-but-insight-poor-how-companies-are-missing-the-point/. Accessed 20 May 2020.

Edelman, D.C. 2010. Branding in the digital age. *Harvard Business Review* 88 (12): 62–69.

Edvardsson, B., and T. Strandvik. 2000. Is a critical incident critical for a customer relationship? *Managing Service Quality* 10 (2): 82–91.

Finne, Å., and C. Grönroos. 2009. Rethinking marketing communication: From integrated marketing communication to relationship communication. *Journal of Marketing Communications* 15 (2–3): 179–195.

Finne, Å., and C. Grönroos. 2017. Communication-in-use: customer-integrated marketingcommunication. *European Journal of Marketing* 51 (3): 445–463.

Gensler, S., F. Völckner, Y. Liu-Thompkins, and C. Wiertz. 2013. Managing brands in the social media environment. *Journal of Interactive Marketing* 27 (4): 242–256.

Grace, D., and O'cass, A. . 2005. Examining the effects of service brand communications on brand evaluation. *Journal of Product & Brand Management*. https://doi.org/10.1108/10610420510592581.

Iacobucci, D., M. Petrescu, A. Krishen, and M. Bendixen. 2019. The state of marketing analytics in research and practice. *Journal of Marketing Analytics* 7 (3): 152–181.

Kamensky, J. 2018. Data rich, but information poor. *IBM Center for The Business of Government*. http://www.businessofgovernment.org/blog/data-rich-information-poor. Accessed 20 May 2020.

Keller, K. L. 2002. Branding and brand equity. *Handbook of Marketing*, 151.

Keller, K.L., and D.R. Lehmann. 2006. Brands and branding: Research findings and future priorities. *Marketing Science* 25 (6): 740–759.

Kelly, D.M., and S. Fairley. 2018. The utility of relationships in the creation and maintenance of an event portfolio. *Marketing Intelligence & Planning* 36 (2): 260–275.

Korfiatis, N., P. Stamolampros, P. Kourouthanassis, and V. Sagiadinos. 2019. Measuring service quality from unstructured data: A topic modeling application on airline passengers' online reviews. *Expert Systems with Applications* 116: 472–486.

Krishen, A.S., and M. Petrescu. 2018. Marketing analytics: Delineating the field while welcoming crossover. *Journal of Marketing Analytics* 6 (4): 117.

Lee, D., K. Hosanagar, and H.S. Nair. 2018. Advertising content and consumer engagement on social media: Evidence from Facebook. *Management Science* 64 (11): 5105–5131.

Lenka, S., V. Parida, and J. Wincent. 2017. Digitalization capabilities as enablers of value co-creation in servitizing firms. *Psychology & Marketing* 34 (1): 92–100.

Lindgreen, A., and I. Crawford. 1999. Implementing, monitoring, and measuring a programme of relationship marketing. *Marketing Intelligence & Planning* 17 (5): 231–239.

Liu, B. 2012. Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies* 5 (1): 1–167.

Lopez, A., E. Guerra, B. Gonzalez, and S. Madero. 2020. Consumer sentiments toward brands: The interaction effect between brand personality and sentiments on electronic word of mouth. *Journal of Marketing Analytics* 8 (4): 203–223.

Maceli, M., 2016. Introduction to text mining with R for information professionals. *Code4Lib Journal*, 33.

Maclaran, P., and M. Catterall. 2002. Researching the social web: Marketing information from virtual communities. *Marketing Intelligence & Planning* 20 (6): 319–326.

Manaman, H.S., S. Jamali, and A. AleAhmad. 2016. Online reputation measurement of companies based on user-generated content in online social networks. *Computers in Human Behavior* 54: 94–100.

Pearson, . 1996. *Building brands directly*. London: MacMillan Business.

Pohl, D., A. Bouchachia, and H. Hellwagner. 2016. Online indexing and clustering of social media data for emergency management. *Neurocomputing* 172: 168–179.

Rambocas, M., and J. Gama. 2013. The role of sentiment analysis. Working Paper [489], FEPUP, University of Porto, April, available at: http://wps.fep.up.pt/wps/wp489.pdf (accessed in November 2020).

Rambocas, M., and B.G. Pacheco. 2018. Online sentiment analysis in marketing research: A review. *Journal of Research in Interactive Marketing* 12 (2): 146–163.

Rosenthal, B., and E.P.Z. Brito. 2017. The brand meaning co-creation on Facebook. *Marketing Intelligence & Planning* 35 (7): 923–936.

Shao, G. 2009. Understanding the appeal of user-generated media: A uses and gratification perspective. *Internet Research* 19 (1): 7–25.

Simon, F., and V. Tossan. 2018. Does brand-consumer social sharing matter? A relational framework of customer engagement to brand-owned social media. *Journal of Business Research* 85: 175–184.

Wang, Z., and X. Ye. 2018. Social media analytics for natural disaster management. *International Journal of Geographical Information Science* 32 (1): 49–72.

Wedel, M., and P.K. Kannan. 2016. Marketing analytics for data-rich environments. *Journal of Marketing* 80 (6): 97–121.

Wilson, A., H. Murphy, and J.C. Fierro. 2012. Hospitality and travel: The nature and implications of user-generated content. *Cornell Hospitality Quarterly* 53 (3): 220–228.

Xu, Z., G.L. Frankwick, and E. Ramirez. 2015. Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research* 69 (5): 1562–1566.

Zhenning Xu is an Assistant Professor of Marketing at California State University, Bakersfield (2019–present). Previously he was an Assistant Professor of Marketing at the University of Southern Maine (2016–2019). His research interests include digital marketing, marketing analytics, research computing, text analytics, and sentiment analysis. He has extensive experience dealing with various marketing problems using data analytics and machine learning with R, Python, Tableau, DataRobot, Perl, Stata, SAS, Google Data Studio, and SPSS. He held administrative positions for about 7 years before pursuing a doctorate in Business. He earned a Ph.D. in Business Administration from the University of Texas at El Paso, a Data Science Specialization (9-course certificate) from Johns Hopkins University (Online), an MBA, and a Bachelor's in Business Statistics. He was the recipient of the 2020 Literati Award Winner for Highly Commended Papers in the Qualitative Market Research. He was also the recipient of the Dodson research grant, the Mike Loya research scholarship, and several other grants. As of 2020, he has published over 15 peer-reviewed papers including the Journal of Business Research, the Journal of Marketing Theory and Practice, the Journal of Macromarketing, Qualitative Market Research, an International Journal, Emerging Markets Finance and Trade, the American Marketing Association (AMS) conference proceedings, the AMS conference proceedings, the INFORMS conference proceedings, the DSI conference proceedings, the IEEE conference proceedings, and the ACM conference proceedings.

Colin Vail is an MBA graduate of the University of Southern Maine with a concentration in Business Analytics. His graduate work and occupation as a Senior Business Analyst covers customer segmentation, big data analytics, and forecast modeling.

Dr. Amarpreet Kohli is an Associate Professor at the School of Business, University of Southern Maine, where he teaches courses on Operations and Supply Chain Management. He received his Ph.D. from University of Louisville, Kentucky. The three primary themes that organize his current scholarly activities are: supply chain collaboration, service operations and text mining/ sentiment analysis. Additionally, he is interested in interdisciplinary research related to diversity, equity, and inclusion issues in higher education. He has published papers in various peer-reviewed journals including Technovation, Supply Chain Forum: An International Journal and International Journal of Information Systems, and Supply Chain Management, and presented at various national and international conferences including DSI, INFORMS, and POMS. Additionally, he has more than ten years of experience in sales and distribution, new product development, and quality assurance in various multinational companies.

Saeed Tajdini received his Ph.D. in Business Administration from the University of Texas El Paso in 2017 and since has been an Assistant Professor of Marketing at Indiana University Southeast. Saeed mainly teaches international business, principles of marketing, and data analytics. His research has been published in the Journal of Marketing Theory and Practice, the Academy of Marketing Science Review, as well as the Qualitative Market Research: An International Journal.