Business Horizons (2018) **61**, 199—210

Available online at www.sciencedirect.com

ScienceDirect

KELLEY SCHOOL OF BUSINESS
INDIANA UNIVERSITY

www.elsevier.com/locate/bushor

Social media analytics for enterprises: Typology, methods, and processes

In Lee

School of Computer Sciences, Western Illinois University, 1 University Circle, Macomb, IL 61455-1390, U.S.A.

KEYWORDS
Social media analytics;
Social media;
Sentiment analysis;
Competitive analysis;
Social network analysis;
Social media intelligence

Abstract This article provides an overview of social media analytics for managers that seek to utilize the practice for social media intelligence. Currently, managers are challenged to analyze an abundance of social media data but lack a framework within which to do so. Toward this end, this article presents a simple typology of social media analytics for enterprises. It also discusses various analytics methods for social media data. Then, this article discusses management processes of social media analytics for enterprises. An illustration of social media analytics is provided with real-world consumer review data. Finally, four challenges are discussed.

1. Social media analytics

Social media analytics refers to the practice of gathering data from social media platforms and analyzing the data to help decision makers address specific problems. Social media analytics have been used by a wide range of people, including social scientists, business managers, and medical professionals. Automated social media analytics is inexpensive and fast compared to traditional media analysis, via which data collection is oftentimes manual and the analysis is labor-intensive. The popularity of social media analytics surged when popular social media platforms allowed enterprises to access enormous amounts of customer data from their sites. Social media platforms focus on idiosyncratic groups of content creators and content consumers. For example, Twitter is a real-time information network that connects users and followers to the latest stories, ideas, opinions, and news. Likewise, Yelp is a site where customers publish unsolicited reviews and viewers read them. According to recent statistics (Mansfield, 2016), Facebook is the most popular social networking site, with over 1.79 billion monthly active users in 2016. Concurrently, the popular photo-sharing site Instagram enjoys 500 million monthly users, who share an average of 95 million photos and videos per day. For its part, the micro-blog hosting site Twitter has roughly 317 million monthly active users and generates 6,000 tweets per second (Mansfield, 2016).

Social media analytics can analyze social media data to obtain consumers' innovative ideas and enhance customer relationships. Many Fortune

E-mail address: i-lee@wiu.edu

0007-6813/$ — see front matter
https://doi.org/10.1016/j.bushor.2017.11.002

500 companies, from McDonald's to Pepsi to Marriott, have employed social media analytics in order to derive a competitive advantage. Multinational hospitality company Marriott International operates its own social media center, named M Live; here, about a dozen employees analyze Twitter feeds, Instagram photos, and Facebook posts in effort to engage clients in social conversations, increase the hotel chain's brand presence, and keep up with the latest trends (Golden & Caruso-Cabrera, 2016). Using a technology called geo-fencing, Marriott is able to monitor every public post on social media platforms made from within company properties . . . and reach out to those customers to prove that Marriott values them. In a similar fashion, the Coca-Cola Company operates a social media center called the Hub Network (Journey Staff, 2014), which deploys a suite of social listening, analysis, and engagement tools to deliver a better, faster, and more efficient response to real-time social media opportunities and issues.

Among many types of organizations, enterprises are the most active users of social media analytics. Analyzing social media data to better understand why customers purchase a product or service plays an important role in sustaining competitive advantage (Brooks, Heffner, & Henderson, 2014). Social media analytics equipped with advanced techniques has significantly affected a company's ability to leverage otherwise unattainable social media intelligence. Enterprises can better understand customer behaviors by combining intelligence acquired by social media platforms with traditional customer intelligence (Sigala & Chalkiti, 2015). According to marketsandmarkets.com (2016), the global social media analytics market will grow from \$1.6 billion in 2015 to \$5.4 billion in 2020 at a compound annual growth rate (CAGR) of 27.6%. The North American region is a leading revenue-generating region for social media analytics vendors, followed by Europe, with a high penetration of social media analytics in multiple industries such as manufacturing, healthcare, transportation, and logistics. Most of the major vendors of social media analytics solutions and services—including SAS, IBM, Oracle, Tableau, and Hootsuite Media—are located in the U.S.

A recent survey of business managers illustrates that social media analytics is viewed as an untapped opportunity, but much must be done to fully exploit social media analytics for consumer packaged goods (CPG) companies (Accenture, 2014). The same survey identifies marketing as the primary area of social media analytics. Companies recognize the value of social media analytics in innovation and product development, followed by customer service, operations, and strategy. Nearly half of the respondents indicated information technology as the department in which analytics competencies and roles are concentrated, followed by an analytic unit within departments and business units, and a centralized analytics unit at the enterprise level.

A variety of open-source tools, commercial toolkits, and proprietary platforms that provide simple standard analytics and customized social media analytics exist for enterprises. Using these tools, innovative managers are finding new ways of automatically collecting, combining, and analyzing data from social media to understand customers, manage customer relationships, and design new products. In our current technology-driven business environment, companies should plan their social media analytics efforts and revise them regularly. However, there exists a scarcity of typologies that may be used by managers to understand types of analytics and identify appropriate methods needed for analyzing content from social media.

In response to the growing interest by enterprises, this article presents a typology of enterprise social media analytics and highlights how its four categories of social media analytics—real-time competitive, non-real-time competitive, real-time customer, and non-real-time customer—can help companies develop social media intelligence. This typology incorporates time and market orientation perspectives and maps a wide range of social media analytics applications into four categories. I then provide a brief overview of social media platforms and discuss important characteristics of these platforms. This article also presents various methods for conducting social media analytics and a four-stage analytics process model. The four-stage analytics process represents the essential stages that facilitate the management of social media analytics. Next, an illustration of social media analytics is provided with the analysis of real-world consumer review data using sentiment analysis and a regression model. Finally, this article discusses the challenges that enterprises face in using social media analytics.

2. Typology of enterprise social media analytics

In this section, I use a typology approach to identify and systematically categorize social media analytics. Broadly speaking, typologies are generated through qualitative classifications rather than quantitative or statistical analyses (Hunt, 1991). While constructing the typology, I took five criteria into consideration: adequacy of phenomenon specification, adequacy of specification of classification characteristic, mutually exclusive categories, collectively exhaustive typology, and usefulness of topology (Hunt, 1991).

As mentioned in many studies, timeliness and market orientation are important in determining the social media analytics adopted by enterprises. Kohli and Jaworski (1990) define *market orientation* as the firm's ability to generate, disseminate, and use superior information about its customers and competitors. This typology employs a market orientation dimension with two dichotomous values (customer and competitor) and a timeliness dimension with two dichotomous values (real-time analysis and non-real-time analysis). Utilizing these dimensions, we present a 2×2 typology of enterprise social media analytics (Table 1) that charts out some of the popular social media analytics used by enterprises. These include real-time customer analytics, non-real-time customer analytics, real-time competitive analytics, and non-real-time competitive analytics.

With the advent of information technologies and social media, social media platforms became more popular than proprietary sources in seeking the information needed to blunt competitive challenges (Harrysson, Metayer, & Sarrazin, 2012). Social media analytics enables companies to rapidly collect data from social media platforms and develop in-depth intelligence regarding the competitive environment (Arrigo, 2014). While the four aforementioned types of social media applications can be used separately, cutting-edge companies will deploy all four simultaneously. Customer analytics and competitive analytics are like two sides of the same coin in that for a company, customers are a target of competitive intelligence on the part of competitors.

Customer analytics is widely supported by commercial service providers; popular metrics include size of social media pages, response time to complaints, number of engagements (comments, likes, shares for each post), and demographics of the people connected to a user. Competitive analytics is a method to gather and analyze data about competitors and the business environment; this may utilize news analytics methods as well as sentiment analysis. Tasks of competitive analysis include:

- *Sentiment tone*: how positive, negative, or neutral the tone of the content is about the competitor.
- *Relevance*: how relevant or substantive the content is for the competitor.
- *Keywords analysis*: what and how many different keywords are used in the content about the competitor.
- *Intensity analysis*: how repetitively the keywords are used over time.
- *Alert analysis*: tracking special announcements such as technology breakthrough, product recalls, change of upper management, and financial performance.

2.1. Real-time customer analytics

Real-time customer analytics allow companies to manage conversations with customers, respond to their issues, and deliver information that is timely and relevant via social media channels in real—or near-real—time (Walters, 2013). Data that languishes without analysis loses its relevance (Qualman, 2009).

Table 1. A typology of enterprise social media analytics

		Timeliness	
		Real-Time	Non-Real-Time
Market Orientation	Customer	Real-Time Customer Social Media Analytics Reactive marketing efforts (e.g., keyword analysis, location analysis, conversation analysis, complaint detection, and alert from online review or comments)	Non-Real-Time Customer Social Media Analytics Proactive marketing efforts (e.g., identification of profitable customer groups, social network analysis, influencer analysis, web analytics, sentiment analysis)
	Competitor	Real-Time Competitive Social Media Analytics Operational intelligence (e.g., monitoring of prices and promotions, news alert, headlines, new product announcement, merger and acquisitions)	Non-Real-Time Competitive Social Media Analytics Strategic and tactical intelligence (e.g., periodic trend analysis of competitors' pricing, new product development, technology development, customer services, complaints, employee comments)

An empirical study shows that the more rapid the response to the consumer, the more valuable that response is (Weiss, Lurie, & MacInnis, 2008). It is important to identify and resolve individualized incidents that may spread quickly via word of mouth. Continuous monitoring of online channels is seen as a practice for preventing personalized issues from damaging the company's reputation. Improved customer intelligence obtained through real-time customer analytics enables the company to formulate a more precise, timely response. For example, Wells Fargo set up a command center to better monitor social media sites such as Twitter, Pinterest, and Facebook; engage with consumers on topics that are trending; and respond quickly to specific customer queries (Wisniewski, 2014). This monitoring is also a way to quickly acknowledge complaints and make an effort to resolve them. As part of its social media outreach, Wells Fargo has a dedicated social care unit to which it refers many of its customer complaints.

2.2. Non-real-time customer analytics

While real-time customer analytics is used to formulate reactive marketing strategies, non-real-time customer analytics focuses on developing proactive marketing strategies based on the analysis of customer data over time. Non-real-time customer analytics may produce simple quantitative metrics such as frequency or duration of webpage visits and demographics and geolocation of visitors, number and growth of members on a social media platform, and conversion rate. In addition to simple metrics, companies collect qualitative information about sentiment or opinion. Using accumulated historical customer comments, companies can analyze how customer behaviors evolve over time and may design marketing messages that have higher reach and response. Toward this end, Starbucks launched its own corporate social networking site, www.mystarbucksidea.com, to inform customers about the company's new products/promotions and analyze customers' needs/expectations about products/services in both non-real-time and real time. Likewise, J.P. Morgan's media relations team actively monitors the financial news discussed on various social media platforms (e.g., Twitter, Snapchat, YouTube), develops social media campaigns, and regularly reports on social media campaign performance and listening insights.

2.3. Real-time competitive analytics

Tracking competitors' business activities in real time and monitoring their online conversations with customers can provide valuable information to strengthen various business activities. To develop real-time competitive analytics effectively, a company should understand how its competitors are utilizing various social media platforms. For example, a company may continuously monitor the frequency and sentiments of customers' postings about a competitor and viewers' reactions on various social media platforms and benchmark them for comparison against their own. Real-time competitive analytics may expand to the analysis of online news about product developments, new features of products, product recalls, cyberattacks, and mergers and acquisitions. Real-time competitive analytics will help a company respond effectively to a competitor's moves or address customers' concerns. For example, restaurants or gyms may need to track competitors' daily online discount deals and develop comparable promotional strategies. Real-time competitive analytics can be deployed to monitor online deal sites such as Groupon and LivingSocial and track the discount rate, duration of the discount offerings, types of discount food or services, and the number of deal purchasers for competitors.

2.4. Non-real-time competitive analytics

While the practice of real-time analytics is valuable for many business activities, not all activities are suitable for real-time analytics. Data collection and processing can become unwieldy when a large number of social media platforms are monitored. For example, a trend analysis is more suitable for non-real time analysis, but consumers' complaints may be monitored in real time. The trend analysis of consumers' responses to competitors' products and services may provide new or improved product ideas. Companies can also investigate the relationship between consumer-generated contents and sales performances. These analyses can be done periodically with a sufficient amount of data for statistical analysis. Lowe's and Home Depot can easily access their best-performing competitor's social media data. They can conduct a trend analysis of customer sentiment, postings on social media platforms, number of followers, demographics and geolocations of customers. The competitors' metrics can serve as a benchmark in developing marketing strategies and improving the home company's products and services.

3. Social media platforms

Social media platforms are online communities via which members seek and share common interests, activities, experiences, and information. Social media platforms provide archival data and real-time feeds as well as sophisticated analytics tools. Since each social media platforms is used differently by individuals, the typology of social media analytics has a clear bearing on the social media platform. For example, Twitter is preferred for real-time dissemination of content, while Facebook and Yelp are preferred for non-real-time content dissemination. Customers actively choose and use social media in response to their specific needs and motivations, and tend to employ a mix of social media platforms. Therefore, managers need to monitor their customers' preferred social media platforms and keep up with advances, as the landscape of the platform market changes constantly and new media rapidly emerge.

Social media platforms enable consumers to be content creators as well as content users (Hajli, 2015). Li, Bernoff, and Groot (2011) use the term *groundswell* to describe a trend in which people use social media to get the information they want from other social media users rather than professional or public sources. Social media represents a variety of dynamic and community-based web applications that place value on the power of distributed knowledge and provide users with rich interactive experiences. There is significant competition among the various social media platforms. For example, Google designed Google+ in an attempt to create its own social media network. LinkedIn is a specialized social networking site focused on professional relationships. Since the mid-2000s, social media has paved new ways of accessing customers for businesses and conducting business activities. Users who interact on social media platforms share their own experiences and activities, engage with each other, and create a stronger sense of community (Gruzd, Wellman, & Takhteyev, 2011; Hays, Page, & Buhalis, 2013).

While there is no universally agreed-upon classification for social media platforms, widely categorized generic platforms include social networking services, blogs, social bookmarking sites, content sharing sites, and opinion sharing sites. Social networking services—such as Facebook, Google+, Snapchat, and LinkedIn—have created massive online communities of people who are willing to network, communicate, and collaborate with each other. Blogs are online journals characterized by short entries and regular updates, typically managed by one person/user that provides interaction with others. WordPress, Tumblr, and Twitter are popular blog hosting sites available to individuals and companies. Social bookmarking is the process of adding, annotating, editing, and sharing bookmarks with others via sites like Delicious, Reddit, Digg, and StumbleUpon. Companies increasingly are using social bookmarking sites to promote new products and services. Content sharing sites enable users to create, store, and share their multimedia files (e.g., photos, videos, audios) with others. Instagram, Pinterest, Flickr, and YouTube are content sharing sites widely used for marketing purposes. Opinion sharing sites enable users to write and publish unsolicited reviews about products, services, merchants, and organizations; they include Glassdoor, Zagat, Yelp, and TripAdvisor. The use of these public social media platforms is widespread across all industries. For example, in the financial industry, major banks such as J.P. Morgan, Morgan Stanley, and Bank of America all are avid users of Facebook, Twitter, Instagram, and Snapchat. It is notable that in addition to public social media platforms such as Facebook and Twitter, companies are creating their own private social media platforms for consumers. Procter & Gamble has established an exclusive website for young girls, beinggirl.com, where visitors can discuss concerns related to being a girl and comment on existing and new P&G products. Beinggirl.com has become a very attractive contribution to Procter & Gamble's marketing campaign.

Many social media sites provide application programming interfaces (APIs) to allow third-party web applications to access social media data. Through APIs, social media sites try to engage and grow developer and user communities. For example, the embedded tweets tool provided by Twitter allows social media analytics tools to access tweets from Twitter. Facebook, Google, Instagram, LinkedIn, Yelp, and many other social media sites provide various APIs to users or organizations. To make social media analytics more effective, companies oftentimes need to simultaneously monitor a number of different social media platforms. For example, companies may monitor Facebook for competitors' marketing efforts, Snapchat for competitors' marketing campaigns to teens and millennials, Twitter for product updates and promotions, Instagram for pictures and short videos of mobile users, and Pinterest for high-quality product pictures and pins. While companies can access and analyze user-generated content via social media APIs, they may also use services provided by commercial analytics service providers that collect social media data and conduct analytics on behalf of client organizations.

4. Methods for social media analytics

As discussed, the aforementioned typology illustrates different ways a company can exploit social media for its business intelligence and decision making. Social media platforms contain a wide range of data types and data volumes. For instance, Yelp primarily collects consumer product review data in text format, while Pinterest mainly collects images and videos. Depending on the social media platform and type of analysis, managers need to choose methods of social media analytics. For example, a company may use sensitivity analysis for consumer review data but may use social network analysis to identify competitors' industry collaborations. In this section, we relate methods for social media analytics to corresponding typology and social media platform(s).

Analyzing social media has become a popular research and business activity due to the availability of user-generated content accessible through web-based APIs as provided by major social media sites (Batrinca & Treleaven, 2015). Analyzing market performance through social media sites has attracted a great deal of attention from researchers. Social media analytics utilizes an array of methods developed to derive specific metrics from social media data. Based on a survey of existing commercial tools, we find that sentiment analysis, social network analysis, and statistical methods are most widely used. Image analysis and video analysis are in early stages of technology development.

4.1. Sentiment analysis

Sentiment analysis, also called opinion mining, refers to the application of computational technologies such as natural language processing and computational linguistics to identify and extract subjective information from vast amounts of user-generated content. Sentiment analysis mainly uses two methods: (1) a machine-learning method with which the analytics learns to evaluate sentiment through acquisition and integration of knowledge gained from a large number of sentiment examples, or (2) a lexical-based method with which the analytics evaluates sentiment by utilizing a large dictionary of prescored words and phrases. Sentiment analysis has been used successfully for businesses activities, including predicting stock market movements, determining market trends, analyzing product defects, and managing crises (Fan & Gordon, 2014). Sentiment analysis needs to be interpreted with a grain of salt, though, due to potential sampling biases in the data (e.g., satisfied customers remain silent while those with more extreme positions express their opinions) (Fan & Gordon, 2014). For example, many online users are lurkers who join an online community with the aim of passively participating for a period of time before making the decision to actively participate in the community (Nielsen, 2006). Lurkers are more likely to post neutral messages that are not detailed enough to give up their identity (Nonnecke, Andrews, & Preece, 2006).

Lexical-based methods use a predefined set of words that carry a specific sentiment. They include simple word or phrase counts; the use of emoticons to detect polarity (i.e., positive and negative emoticons used in a message) (Park, Barash, Fink, & Cha, 2013); sentiment lexicons (i.e., based on words in the lexicon that have received specific features marking positive or negative terms in a message) (Gayo-Avello, 2011); and the use of psychometric scales to identify mood-based sentiments. Machine-learning methods often rely on the use of supervised and unsupervised machine-learning techniques. One advantage of machine-learning methods is their ability to adapt and create trained models for specific purposes and contexts. On the other hand, it is known that labeling data might be costly or even prohibitive for some tasks. Sentiment analysis divides into the following specific subtasks (Batrinca & Treleaven, 2015):

- *Sentiment context*: to extract sentiment, one needs to know the context of the contents, which may vary significantly from specialty sites to general sites where contents can cover a wide spectrum of topics.
- *Sentiment level*: text analytics can be conducted at the word, sentence, or document level.
- *Sentiment orientation/polarity*: the sentiment in a text can be positive, neutral, or negative.
- *Sentiment strength*: the strength of a sentiment in a text can be weak, moderate, or strong.
- *Sentiment subjectivity*: a given text may be an opinion or a fact.

4.2. Social network analysis

Social network analysis is the process of analyzing structures of social networks based on *social network theory,* which seeks to explain how networks operate and analyze the complex set of relationships within a network of individuals or organizations (Scott, 2012; Wasserman & Faust, 1994). In a social network, *nodes* are the individual actors and *ties* are the relationships between the actors. Social network analysis provides both visual and mathematical analyses of actor relationships within a network by modeling social network dynamics and growth (e.g., network density, network centrality, network flows). Actors and social ties are important in information dissemination and propagation, including both weak and strong ties (Brown & Reingen, 1987; Datta, Chowdhury, & Chakraborty, 2005). Weimann (1983) found that information becomes influential due to strong ties within the group, ties which are more likely to be used and perceived as credible than weak ties. Social network analysis uses a variety of techniques pertinent to understanding the structure of the network (Scott, 2012). These range from simpler methods such as counting the number of edges a node has or computing path lengths, to more sophisticated methods that compute eigenvectors to determine key nodes in a network (Fan & Gordon, 2014).

Social networking sites such as Facebook, Twitter, and LinkedIn have provided fertile ground for advancing online social network theories and practices. Social networking sites allow the establishment of links connecting family members, friends, and peers (Tuten, 2008). Social networking sites provide a central point of access and bring structure to the process of personal information sharing and online socialization (Jamali & Abolhassani, 2006). Understanding the dynamics of interactions between users can assist in identifying influencers to target in branding and ad campaigns (Chen, Wang, & Yang, 2009). Well-connected users are particularly important for social networking sites, as these users can be highly relevant for the promotion of brands, products, and viral marketing campaigns (de Valck, van Bruggen, & Wierenga, 2009).

4.3. Statistical analysis

Traditional statistical methods have also been used for advanced analytics. Some of the statistical methods include Markov chain Monte Carlo methods, regression models, logistic regression, factor analysis, and cluster analysis. These statistical methods typically require transformation of the original contents into a coded format suitable for statistical methods. Regression analysis has been useful for understanding the causal relationship between various factors obtained from factor analysis. Xiang, Schwartz, Gerdes, and Uysal (2015) applied a text analytical and statistical approach to a large quantity of consumer reviews extracted from Expedia.com to deconstruct hotel guest experiences and examine its association with satisfaction ratings. First, the authors conducted a factor analysis to identify the underlying factors of customer reviews. Then, they performed a linear regression analysis to examine the relationship between guest experience and satisfaction using the identified factors as independent variables and average satisfaction rating as the dependent variable. Similarly, Qu, Zhang, and Li (2008) analyzed the content of review data collected from Yahoo's merchant review system. Based on the content analysis, these authors selected 14 factors and conducted regression analysis to identify factors that have major impacts on customers' evaluation of online merchants.

4.4. Image and video analysis

Images and videos are also significant components of social media data due to rapidly growing content sharing platforms such as YouTube, Instagram, and Flickr. Billions of images are uploaded every day, and mining these images can provide other types of insight beyond what can be captured solely from text. While the aforementioned methods have been developed primarily for text mining, analytics for images and videos have thus far been marginally exploited. Image analysis is the process of organizing images into an archive that is fully searchable and analyzable. Basic image analysis involves the statistical analysis of tag data, demographic data, and download frequency (e.g., Instagram account's average engagement per photo, keyword analysis for comments, most active followers, top locations). Advanced image analysis utilizes image processing techniques, image recognition, and image tags. Image analysis enables companies to mine image data to extract valuable information such as location of people and fashion trends. Similar to basic image analysis, basic video analysis typically involves quantitative metrics such as number of users, response rate, subject, and location. More advanced techniques include accessing video clips posted to social media sites, analyzing voice to determine emotional state of the user, and applying a behavioral model to spoken words to determine personality type of the user.

5. Processes for social media analytics

This section presents a four-stage analytics process that represents a minimum set of stages that facilitate the management of social media analytics. Stage 1 focuses on developing key performance/evaluation metrics. Stage 2 focuses on choosing social media platforms that generate data and monitoring/listening to the chosen social media platforms. Stage 3 focuses on analyzing data using various analytics tools. Stage 4 involves building social media intelligence. The processes are iterative and evolve over time as new social media platforms and tools emerge and the environment changes.

5.1. Stage 1: Develop key social media metrics

Extracting and delivering value from social media data requires precise measurement, so selecting the right metrics is important in connecting data with desired outcomes (Grimes, 2013). Companies must decide the objectives of social media analytics. Some of the widely-used objectives include building brand awareness, increasing website visits, increasing revenue, enhancing corporate reputation, and improving customer service/customer satisfaction. These objectives must be accompanied by both quantitative metrics (e.g., number of followers or fans, number and frequency of postings, number of likes or tweets) and qualitative metrics (e.g., sentiment orientation, strength, subjectivity) necessary to understand the attitudes and opinions of customers and competitors. Regarding competitive analytics, the company must identify competitors to benchmark.

Notably, metrics can guide the choice of methods of social media analytics and platforms. For example, companies can measure size of the community (real-time analysis) and whether it is growing (non-real-time analysis) in Facebook (platform selected). They can use statistical methods to model growth of their social media platforms and find demographics of users. Popular metrics include proportion of active member users, time spent on social media sites, quantity of discussion about the company and its products, redistribution of content to other social media platforms, and improvement of reputation/image of the company.

5.2. Stage 2: Choose, monitor, and listen to social media platforms

Specific information obtained from engaging in social listening activities is highly dependent on where we listen (e.g., selectively monitoring a single social media site vs. monitoring multiple sites) and how the data is analyzed (Schweidel & Moe, 2012). Selection of social media platforms depends upon the impact of those platforms on the business (e.g., number of customers that use them) and accessibility of content on those social media platforms. Forrester Research (2011) defines *social listening* as "programs that utilize social media to monitor, measure, and respond to customer conversations and feedback online." Social media monitoring offers organizations quick access to valuable information about users' consumer profiles, brand awareness/interest/liking/preference, and user ability to understand the brand (Lin & Rauschnabel, 2016). Monitoring customers' conversations on social media sites allows companies to respond quickly to both positive and negative feedback, provide timely customer service, and prevent reputation damage. Companies that monitor and listen to social media conversations will also have the opportunity to counter competitors' moves in quick succession.

Social media listening connects firms to all available market information, and allows them to potentially achieve greater consumer satisfaction and better business performance (Bose, 2008). Free social media tools also exist to track social media in real time such as Google Alerts, Google Trends, Social Mention, and Twitter Search. By continuously listening to social media conversations, companies can acquire data about customers' perceptions about the brand at any time, greatly improve customer service, and create a strong online community.

5.3. Stage 3: Perform social media analytics

The ability of decision makers to collect, filter, and interpret data, messages, and signals has a critical bearing on their strategy (Makadok & Barney, 2001). Social media analytics aim to monitor, filter, and analyze the discussions that take place on social media platforms, providing a comprehensive picture of consumer opinions regarding products and services. Sentiment analysis can suggest whether a product, service, or customer support is better or worse than industry average. Correlating sentiment to recent changes in product design, for example, could provide essential feedback. A number of analytical tools and methods are available to help managers conduct social media analytics. While some social media data are structured, social media data are mainly unstructured and semistructured, leading to high diversity, ambiguity, and textual disorder. These unstructured and semistructured data require preprocessing and scrubbing before data analytics to eliminate missing data, incorrect data, and inconsistent data. The process of data cleaning may involve spell checking, removal of typographical errors or duplicates, validating and correcting values against a known list of entities, and tagging data with metadata. It is also necessary to transform the cleaned data into structured data.

5.4. Stage 4: Build social media intelligence

Social media analytics retrospectively describes what has already happened. It does not prescribe or guide an organization's next steps (Moe & Schweidel, 2014). *Social media intelligence* enables managers to prescribe what should be done with the results of social media analytics. Social media intelligence is achieved by combining knowledge generated from traditional intelligence activities and knowledge gained from social media analytics, and helps managers develop better actionable market decisions that align with the company's objectives (Walters, 2013). According to Moe and Schweidel (2014), social media intelligence allows the company to:

- Understand the behaviors driving the creation of online opinions from both a psychological and sociological perspective;
- Assess the implications of these behaviors on how we interpret social media; and
- Integrate these insights into an overall strategy.

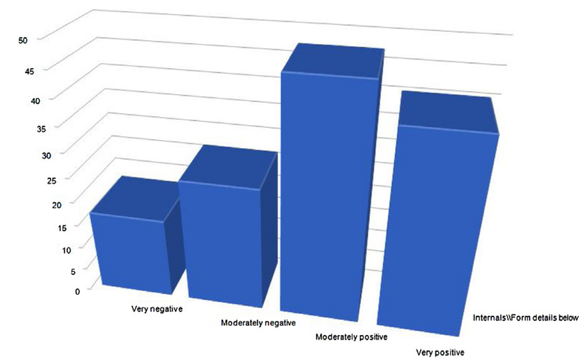
To achieve these objectives, the activities of social media intelligence need to be comprehensive and interdisciplinary, and include social monitoring and listening, social media analytics, social media platform management, data management, and sharing of analytic outcomes with stakeholders. To manage social media intelligence activities effectively, firms such as GM, Marriott International, and Walmart are increasingly establishing centralized social media centers to coordinate various social media activities that span multiple disciplines and departments.

6. An illustration: Customer analytics using merchant review data

This section illustrates the use of customer analytics with real merchant review data. A large restaurant in a suburban Chicago area was selected for the analysis of customer review data collected from Yelp, one of the most popular opinion sharing sites; I used NVivo analytics software. For the restaurant, 50 customer postings were collected. Each post consisted of one comment and one review score. I conducted two-stage analytics. First, sentiment analysis was used to categorize 130 sentences of the 50 customer postings into sentiment polarity: very positive (40 sentences), moderately positive (47 sentences), moderately negative (26 sentences), and very negative (17 sentences). Figure 1 shows visualized sentiment categories.

Then, the 130 sentences were further categorized into six factors: price, service, food, atmosphere, condition (e.g., restaurant cleanliness, accessibility), and selection (e.g., menu variety). These six factors served as independent variables and the review score as a dependent variable. Note that each comment can contain multiple factors. A coding was conducted manually to convert sentiment polarity into specific numeric codes for the factors. A 'very positive' sentence was assigned +4, 'moderately positive' sentence assigned +2, 'moderately negative' assigned −2, and 'very negative' assigned −4. Through the coding process, 50 original customer postings were converted into a sample size of 50 for multiple regression analysis. The seven variables and their descriptive statistics are shown in Table 2.

Figure 1. Visualized sentiment categories



As such, our regression model of the form

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

is

$$Y = 2.91 + 0.256X_1 + 0.168X_2 + 0.478X_3$$

where Y is the review score. This model has an R-square value of 0.664.

A stepwise regression was run on these seven variables. Table 3 shows the beta coefficients of these variables as well as their p-values. The regression was run with a 5% level of significance. The analysis shows that the review score for the restaurant has a high correlation with (1) food, (2) price,

Table 2. Descriptive statistics

	Mean*	Standard Deviation	n
Price	−0.76	1.975	50
Service	0.08	1.066	50
Food	0.92	2.625	50
Atmosphere	−0.08	0.566	50
Condition	0.12	0.627	50
Selection	0.16	0.681	50
Review Score	2.98	1.152	50

* The midpoint is zero

Table 3. Significant coefficients

Variable	Unstandardized Beta	Standardized Beta	*p-value*
β_0 - Constant	2.910		.000
β_1 - Food	.256	.588	.000
β_2 - Price	.168	.288	.007
β_2 - Atmosphere	.478	.235	.009

and (3) atmosphere. As expected, food was the most important factor for customers' overall satisfaction, following by price and atmosphere. However, service, condition, and selection were not significant factors.

This section illustrates that a powerful social media analytics model can be developed with a manageable amount of data for small and medium-sized enterprises to discover relationships between consumers' comments and review scores. While more comprehensive social media analytics might add more value to merchants, this illustration shows that even simple social media analytics can provide considerable marketing ideas to enterprises. Furthermore, the merchant can also analyze its competitors' review data to gain insights into business operations and make marketing decisions accordingly.

7. Challenges

Certain challenges are associated with enterprise adoption of social media analytics. This section identifies four challenges: bias in social media data, selection of social media metrics, noise in social media data, and unstructured social media data. These challenges must be taken seriously and mitigated by enterprises to take full advantage of social media analytics.

7.1. Bias in social media data

Social media platforms are able to reach only a subset of customers. Therefore, social media data may suffer from an unrepresentative customer population that prevents the data from being extrapolated to the general public. While some participants prefer to be active by posting comments and creating content on social media, others prefer to be reticent (Miller & Lammas, 2010). Rather than write something false, some consumers may share only part of what they do or practice in real life to portray a certain status, and thus information may be only partially accurate or complete (Moe & Schweidel, 2014). Fake reviews are oftentimes posted by unscrupulous companies, zealous competitors, disgruntled employees, and unhappy consumers. While managers may be aware of these issues and make efforts to mitigate the effects of such biases in social media data, they need to develop new methods to correct this in the first place.

7.2. Selection of good social media metrics

Social media platforms allow for the easy collection of popular metrics such as number of views, click-through rates, and frequency of posting opinions on blogs. It is notable that there is a diminishing return on investment (ROI) in social media analytics, but oftentimes managers collect data for the sake of measuring or due to easiness of measuring. It is imperative to identify and prioritize key metrics, establish target values, and generate periodic performance reports. Selecting good social media metrics is challenging due to lack of expertise and experience. While metrics may measure increase in revenue or decrease in expense in the long term, they may not necessarily reveal changes in customers' behavior immediately. As such, a time gap should be taken into account when measuring the ROI of social media analytics. Considering social media analytics is a new discipline, managers should also see if advances in social media analytics make collecting new metrics worth the investment.

7.3. Noise in social media data

Sifting through noise in data is challenging. Managers need to recognize the possibility of noise being mixed in among messages of value. There are many sources of noise in social media data, including spam, fake reviews, false accounts, and duplicate content. Filtering and scrubbing data are important for the quality of social media analytics. However, current social media analytics offer limited automated filtering and scrubbing capabilities. Automated procedures, combined with manual procedures, may mitigate the noise of the data. Researchers need to develop comprehensive data filtering and scrubbing methods. Visualization and artificial intelligence represent promising methods of noise reduction.

7.4. Unstructured social media data

Most social media content—including email, news articles, letters, images, audio, and videos—is unstructured. Analysis of unstructured data can help companies gain valuable insights into their customers, products, services, and competitors. Relational database management systems (RDBMS) developed to manage structured data are not suitable for managing unstructured data. Real-time analytics can be more challenging in integrating and analyzing unstructured data from multiple sources due to noise in the data and different semantics used for the same things. While numerous tools are available for companies to analyze unstructured data, managers face issues regarding poor data quality, integrating structured and unstructured data, and massive data volume. They need to understand the type and value of information gained from unstructured data and select the right analytics and data management tools to extract maximum value from unstructured data. It is also important to take note of advances in the data management technologies for unstructured data such as Hadoop and NoSQL.

8. Conclusion: Finding ways to analyze and compete

As social media was widely adopted by customers, it became imperative for businesses to leverage social media platforms to stay competitive in the global economy. Social media analytics are used to monitor and listen to word-of-mouth that spreads in social media platforms, and conduct thorough analyses of consumer opinions on products and services. With the growth of social media platforms, the influx of big data became a huge challenge (O'Reilly & Lancendorfer, 2014). Therefore, it is crucial that companies select the right social media platforms and the correct types of data to collect and analyze.

This article presented a 2×2 typology of enterprise social media analytics that charts out some of the popular social media analytics used by enterprises: real-time customer analytics, non-real-time customer analytics, real-time competitive analytics, and non-real-time competitive analytics. Then, this article discussed various social analytics methods including sentiment analysis, social network analysis, statistical methods, and image and video analytics. An illustration of social media analytics with real data involving sentiment analysis and a regression model was provided for managers' understanding of this important area. Finally, four challenges were identified.

To gain competitive advantage, enterprises need to monitor and analyze not only their customer-generated content but also competitors' customer-generated content on various social media sites. While our discussion here is limited to direct competitors, social media analytics can be expanded to the analysis of supply chain to understand current standing on customer relationships and competitiveness, and improve supply chain efficiency and effectiveness. It is also worth mentioning that while this article focuses on social media analytics, social media analytics is part of large business data analytics, which involves non-social media data for overall business intelligence.

Acknowledgment

I offer special thanks to Editor Dr. Dan Li for her valuable comments and suggestions.

References

Accenture (2014). *What's trending in analytics for the consumer packaged goods industry?* Dublin, Ireland: Accenture.

Arrigo, E. (2014). Social media opportunities for market-driven firms. In I. Lee (Ed.), *Integrating social media into business practice, applications, management, and models* (pp. 180–199). Hershey, PA: IGI Global.

Batrinca, B., & Treleaven, P. C. (2015). Social media analytics: A survey of techniques, tools, and platforms. *AI & Society, 30*(1), 89–116.

Bose, R. (2008). Competitive intelligence process and tools for intelligence analysis. *Industrial Management & Data Systems, 108*(4), 510–528.

Brooks, G., Heffner, A., & Henderson, D. (2014). A SWOT analysis of competitive knowledge from social media for a small start-up business. *Review of Business Information Systems, 8*(1), 23–34.

Brown, J. J., & Reingen, P. H. (1987). Social ties and word-of-mouth referral behavior. *Journal of Consumer Research, 14*(3), 350–362.

Chen, W., Wang, Y., & Yang, S. (2009). Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 199–208). New York, NY: ACM.

Datta, P. R., Chowdhury, D. N., & Chakraborty, B. R. (2005). Viral marketing: New form of word-of-mouth through internet. *Business Review: Federal Reserve Bank of Philadelphia, 3* (2), 69–75.

de Valck, K., van Bruggen, G. H., & Wierenga, B. (2009). Virtual communities: A marketing perspective. *Decision Support Systems, 47*(3), 185–203.

Fan, W., & Gordon, M. D. (2014). Unveiling the power of social media analytics. *Communications of the ACM, 57*(6), 74–81.

Forrester Research. (2011). *Listening and engaging in the digital marketing age*. Available at http://i.dell.com/sites/doccontent/corporate/secure/en/Documents/listening-and-engaging-in-the-digital-marketing-age.pdf

Gayo-Avello, D. (2011). Don't turn social media into another 'Literary Digest' poll. *Communications of the ACM, 54*(10), 121—128.

Golden, J., & Caruso-Cabrera, M. (2016, August 3). Why Marriott is so interested in your social media. *CNBC*. Available at http://www.cnbc.com/2016/08/02/why-marriott-looks at-what-you-post-on-social-media-from-your-room.html

Grimes, S. (2013). The rise and stall of social media listening. *Information Week, 1361*, 5—6.

Gruzd, A., Wellman, B., & Takhteyev, Y. (2011). Imagining Twitter as an imagined community. *The American Behavioral Scientist, 55*(10), 1294—1318.

Hajli, N. (2015). Social commerce constructs and consumer's intention to buy. *International Journal of Information Management, 35*(2), 183—191.

Harrysson, M., Metayer, E., & Sarrazin, H. (2012). How 'social intelligence' can guide decisions. *McKinsey Quarterly, November,* 1—9.

Hays, S., Page, S. J., & Buhalis, D. (2013). Social media as a destination marketing tool: Its use by national tourism organizations. *Current Issues in Tourism, 16*(3), 211—239.

Hunt, S. D. (1991). *Modern marketing theory: Critical issues in the philosophy of marketing science*. Cincinnati, OH: South-Western Publishing Co.

Jamali, M., & Abolhassani, H. (2006). Different aspects of social network analysis. In *IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 66—72). Hong Kong, China: IEEE.

Journey Staff. (2014, November 3). The Hub Network fuels Coke's real-time marketing. *The Coca-Cola Company.* Available at http://www.coca-colacompany.com/stories/the-hub-network-fuels-cokes-real-time-marketing

Kohli, A. K., & Jaworski, B. J. (1990). Market orientation: The construct, research propositions, and managerial implications. *Journal of Marketing, 54*(2), 1—18.

Li, C., Bernoff, J., & Groot, M. (2011). *Groundswell*. Boston, MA: Harvard Business Review Press.

Lin, C. A., & Rauschnabel, P. A. (2016). Social media marketing: Psychological insights, managerial implications, and future research directions. In I. Lee (Ed.), *Encyclopedia of e-commerce development, implementation, and management* (pp. 2144—2158). Hershey, PA: IGI Global.

Makadok, R., & Barney, J. B. (2001). Strategic factor market intelligence: An application of information economics to strategy formulation and competitor intelligence. *Management Science, 47*(12), 1621—1638.

Mansfield, M. (2016). Social media statistics 2016. *Small Business Trends*. Available at https://smallbiztrends.com/2016/11/social-media-statistics-2016.html

Marketsandmarkets.com. (2016). *Social media analytics market worth 5.40 billion USD by 2020*. Available at http://www.marketsandmarkets.com/PressReleases/social-media-analytics.asp

Miller, R., & Lammas, N. (2010). Social media and its implications for viral marketing. *Asia Pacific Public Relations Journal, 11*(1), 1—9.

Moe, W., & Schweidel, D. A. (2014). *Social media intelligence*. Cambridge, UK: Cambridge University Press.

Nielsen, J. (2006). The 90-9-1 Rule for participation inequality in social media and online communities. *Nielsen Norman Group*. Available at http://www.nngroup.com/articles/participation-inequality

Nonnecke, B., Andrews, D., & Preece, J. (2006). Non-public and public online community participation: Needs, attitudes, and behavior. *Electronic Commerce Research, 6*(1), 7—20.

O'Reilly, K., & Lancendorfer, K. M. (2014). Using the power of social media marketing to build consumer-based brand equity. In I. Lee (Ed.), *Integrating social media into business practice, applications, management, and models* (pp. 56—77). Hershey, PA: IGI Global.

Park, J., Barash, V., Fink, C., & Cha, M. (2013). Emoticon style: Interpreting differences in emoticons across cultures. In *International AAAI Conference on Weblogs and Social Media* (pp. 466—475). Palo Alto, CA: AAAI.

Qu, Z., Zhang, H., & Li, H. (2008). Determinants of online merchant rating: Content analysis of consumer comments about Yahoo merchants. *Decision Support Systems, 46*(1), 440—449.

Qualman, E. (2009). *Socialnomics: How social media transforms the way we live and do business*. Hoboken, NJ: John Wiley & Sons.

Schweidel, D. A., & Moe, W. W. (2012, October). *The perils of 'selective listening' in social media monitoring: Sentiment and venue choice in social media posting behavior.* DOI: 10.2139/ssrn.1874892

Scott, J. (2012). *Social network analysis*. Thousand Oaks, CA: Sage.

Sigala, M., & Chalkiti, K. (2015). Knowledge management social media, and employee creativity. *International Journal of Hospitality Management, 45*, 44—58.

Tuten, T. L. (2008). *Advertising 2.0: Social media marketing in a Web 2.0 world*. Westport, CT: Praeger.

Walters, S. (2013). Beyond listening: Six steps for integrating and acting on social media. *Business Intelligence Journal, 18*(1), 13—17.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, UK: Cambridge University Press.

Weimann, G. (1983). The strength of weak conversational ties in the flow of information and influence. *Social Networks, 5*(3), 245—267.

Weiss, A. M., Lurie, N. H., & MacInnis, D. J. (2008). Listening to strangers: Whose responses are valuable, how valuable are they, and why? *Journal of Marketing Research, 45*(4), 425—436.

Wisniewski, M. (2014, March 27). Wells Fargo sets up war room to monitor social media sites. *American Banker*. Available at https://www.americanbanker.com/news/wells-fargo-sets-up-war-room-to-monitor-social-media-sites

Xiang, Z., Schwartz, Z., Gerdes, J. H., Jr., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management, 44*, 120—130.