



Sentiment Analysis of Consumer Base with the Utilization of Twitter

Nicholas Oram, Graduate Student, Mercyhurst University

INTRODUCTION

Social media platforms have transformed the customer from being a passive consumer of business information to becoming an active participant in the creation and sharing of business information. Recent data has demonstrated that nearly 30% of online users aged 18 or older participate in various forms of self-created content sharing, with an even higher percentage for those who post comments on a multitude of websites (Trainor 2012). A recent report by the Mintel Group found that 41% of online adults are engaged with their favorite brands through social media applications (Trainor 2012). Forrester Research conducted a survey that found that 37% of online U.S. adults post product reviews, comment on blogs, or add information to online forums. Moreover, 21% of those surveyed posted their own original information onto social media websites and 69% only read information on various social media platforms (Baker 2009).

Customer involvement through social media is a necessity for any business since engaged customers will play an essential role in viral marketing campaigns. (Agnihorti et al. 2012). Viral marketing is defined as: “electronic word of mouth whereby some form of marketing message related to a company, brand, or product is transmitted in an exponentially growing way (Kaplan & Haenlein, p. 253, 2011).” Therefore, proper

social media analysis will allow viral marketing to reach an increasingly growing consumer clientele and significantly increase customer engagement. Customer engagement is fulfilled when constant interaction between customer and company is created that will allow for relationships that create mutual loyalty between both parties (Agnihotri et al., 2012).

Microblogging is defined as, “a form of blogging that lets you write brief text updates (usually less than 200 characters) about your life on the go and send them to friends and interested observers via text messaging, instant messaging (IM), email, or the web (Java et al., 2007, p. 1).” The growth of microblogging, especially with the use of Twitter, will make it necessary for companies to analyze both their customers’ conversations over microblogs, but most significantly utilize pull marketing tactics to drive consumer content. As compared to regular blogging, microblogging allows the consumer to increase the speed of their communications. The shorter time period decreases the users’ requirement of time to post comments about any company they follow and allows the user to post multiple updates a day in a matter of seconds (Java et al., 2007). Most significantly, (Jansen et al., 2009) Jansen reported that 20% of microblog postings contained brand sentiment, with 50% of these postings demonstrating positive sentiment with 33 percent of microblog posts being critical of a company.

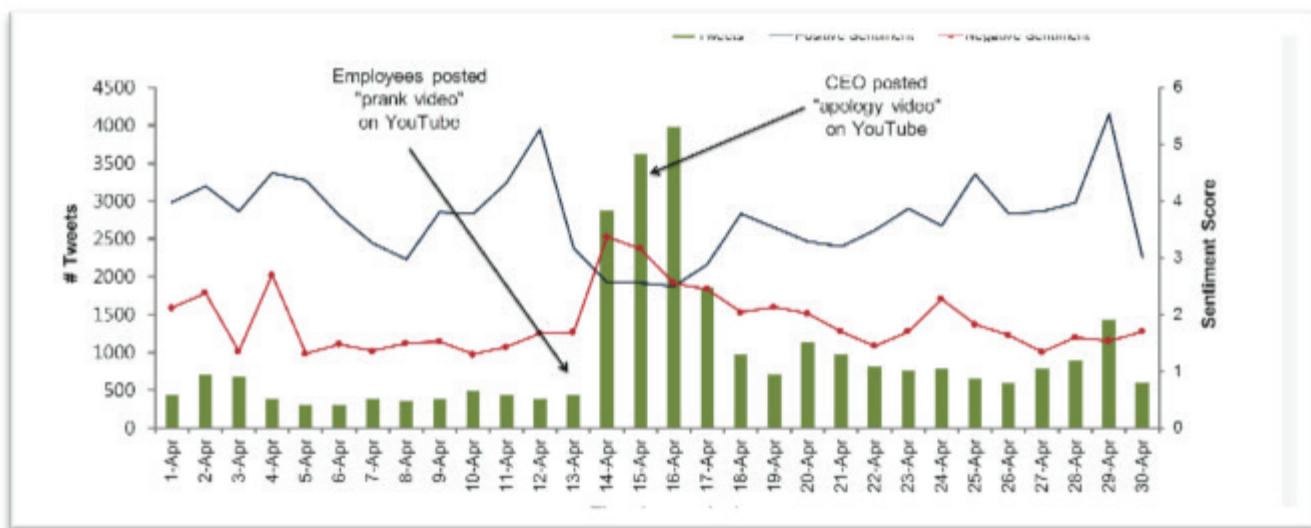


Figure 1: Sentiment analysis of Domino's Pizza
Source: Park et al., 2012

One such example of the benefits of monitoring microblogging activity occurred in 2009 when two Domino's Pizza employees filmed a prank in the restaurant's kitchen while making customers' food. The employees posted the video onto YouTube and the video quickly spread through the use of Twitter, with more than half a million views within two days. With the videos unveiling, customers conveyed tremendous negative sentiment, making it necessary for President of Domino's Pizza Patrick Doyle to issue an apology over YouTube (Figure 1). Additionally, Domino's Pizza had to create a Twitter account for customers to address their grievances with the video. Even though Domino's Pizza was able to address the negative sentiment with the use of social media applications, it is clear that the situation could have quickly worsened over a longer duration of time. Negative sentiment could have significantly damaged the brand reputation of the company (Park et al., 2012).

VALUE OF SOCIAL MEDIA

Bughin et al., (2012) analyzed four industry sectors that represent 20% of the global industry (Financial Services, Consumer Packaged Goods, Professional Services, & Advanced Manufacturing) that demonstrated an increase in revenue potential of \$900 billion to \$1.3 trillion with the utilization of social media platforms. Two-thirds of the potential value in these four industry

sectors will be created by improved communication and collaboration across consumer and business entities (Bughin et al., 2012). A study conducted by Siemens Enterprise Communications found that 70% of consumers want to interact with businesses with the use of social media (Howell, 2010). The problem that has arisen is that only 30% of companies are equipped to handle social media analysis or understand the proper procedures to utilize social media analysis correctly (Howell, 2010).

Driven with the pressures to move towards social media technologies for analysis and marketing endeavors, companies are increasing their expenditures to reach and analyze their consumer base. Forrester Research expects social media expenditures to expand in the U.S. at a rate at 34% from 2010-2014 and to reach a net spending of USD \$3.1 billion. Social media marketing and analysis spending is expected to more than double the growth in spending of all other online marketing strategies by 2014 (Larson & Watson, 2011).

Social media marketing does not seek to replace traditional forms of marketing communications with consumers, but it does seek to complement them in order facilitate more efficient and easier communication methodologies. The new relationship between customer and company has been termed social customer relationship management (SCRM) (Andzulis et al. 2012). However, it is important to clarify what the traditional form of customer relationship management (CRM) is and demonstrate how SCRM adds to CRM's value creation

processes. CRM has been defined as, “The management of the dual creation of value, the intelligence use of data and technology, the acquisition of customer knowledge, and the diffusion of this knowledge to the appropriate stakeholders (Trainor, p. 318, 2012).”

SCRM is similar to the traditional form of CRM, but SCRM seeks to incorporate the newly created technological and social changes that have originated from the introduction of social media platforms into the processes that create and maintain customer value. SCRM is not a replacement for the traditional form of CRM, but should be viewed as an extension that has the capability to add social functions and processes that will have the proficiency to incorporate the interaction between customers, peers, and businesses (Trainor 2012). Figure 2 Displays the intra-group communication with the SCRM model.

TRANSITION OF SOCIAL MEDIA ANALYSIS ENDEAVORS

The popularity of social media analysis and advertising within the business environment began with the use of Facebook. The use of Facebook for marketing

has the goal of reaching two potential audiences: those individuals who have liked the brand and are considered a fan, and those individuals who are friends with the fan who liked the company's Facebook page. Overall, the value of a fan is measured with the following metrics: increasing the depth and loyalty among fans, generating an incremental purchasing behavior, and leveraging the ability to influence the friends of their current fans to spread the company's social media marketing of their various products and services (Lipsman et al., 2012).

The manner in which ROI was originally calculated for the value of a fan over Facebook was measured by the increase in the amount of likes on the company's official Facebook page, how many people shared the company post, and the amount of comments that were generated on a particular post utilized for marketing endeavors (Narayanan et al., 2012). Moreover, companies utilize Twitter in a similar fashion to measure their ROI of a fan with measuring the engagement of their consumer base, which includes: tracking user retweeting rate, audience retweeting rate, and the amount of followers on the company's official Twitter account (Narayanan et al., 2012).

Ballave (2013) affirms that companies have become too heavily reliant on metrics that include measuring

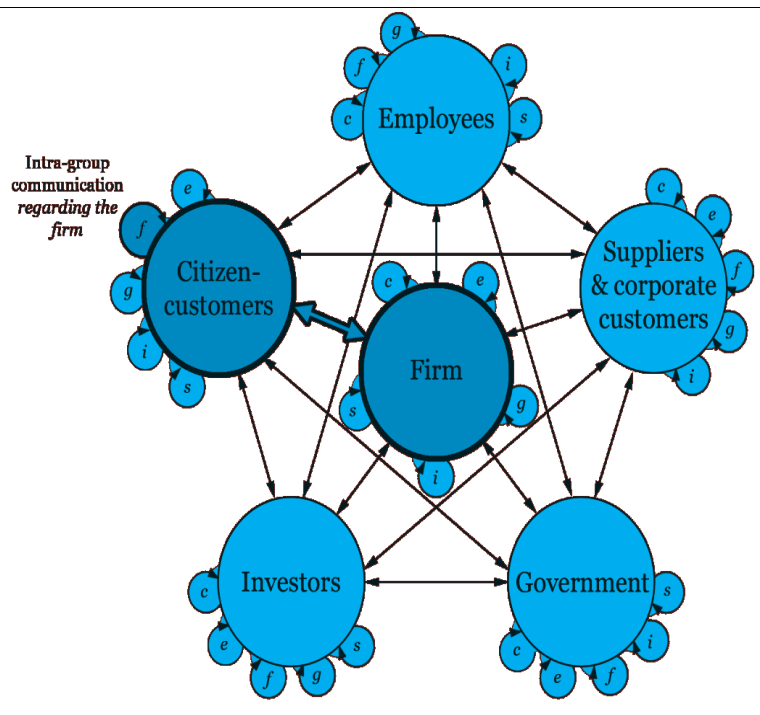


Figure 2: Social media interaction ecosystem

Source: Larson & Watson, 2011

Facebook likes and followers as a successful manner to measure their earned media campaigns. The problem with this type of social media marketing initiative is that it focuses only on building an audience, not participating directly into their consumers' conversations online. Director of Marketing for the startup Uberflip, Neil Bhapkar states that web analytics, especially through social media applications, have gone far beyond key performance indicators (KPI's) of tracking hits on a site such as likes or followers. The most crucial KPIs to measure ROI with social media marketing are unique visits, geography of followers, mobile readership, bounce rates/time spent on the site, heat maps and click patterns, page views, and, most significantly, analyzing consumer sentiment over social media applications (Bhapkar, 2013). Conducting consumer sentiment with Twitter will be described in three different techniques, which include topic based sentiment, stream analysis, and high-density geo-mapping of consumer sentiment.

SENTIMENT ANALYSIS AND TWITTER

The use of Twitter to convey consumers' opinions about their company's products and services is becoming commonplace within the business environment. As a microblogging service, Twitter allows the customer to post in a matter of seconds their feelings towards their favorite products and services, or to share an experience another consumer had with a company. Thus, Twitter has become the most efficient manner in which companies can gather consumer opinions towards their brand and possibly even gain insight into what their competitor's consumer base is communicating via the Twitter platform (Zhang et al., 2011).

Twitter is analyzed according to a variety of methods in order to conduct general sentiment. Emoticons, which are pictorial representations of facial expressions, convey the user's mood and communicate general sentiment about the user feelings within their designated tweet. Hashtags (#) mark general topics and usually are put within a tweet to become visible to a larger audience over Twitter. The special symbol of a re-tweet (RT) indicates when the users' tweet is shared among Twitter. Another important metric is when a Twitter user is targeted (@), meaning they were mentioned within the tweet. Measuring users who have been targeted within a tweet can display what Twitter followers are most influential. These Twitter users could possibly be utilized by a company as a brand ambassador based on their influence over Twitter (Kumar & Sebastian 2012). Sentiment analysis can be conducted

at a variety of different levels, which include: word level, phrase or sentence level, document level, and feature level. However, in terms of its applicability to Twitter, the word and phrase/sentence level is the most significant to judge sentiment of a company's consumer base (Kumar & Sebastian, 2012).

To be able to manipulate and gather the high volume of data that is available through Twitter, companies can utilize three sentiment analysis visual techniques, which include:

- Topic-based sentiment analysis that extracts, maps, and measures customer opinions (explained in personal application section)
- Stream analysis that identifies tweets based on their density
- Pixel-cell based sentiment calendars and high-density geo-mapping that has the ability to visualize a large amount of Twitter data (Hao et al., 2011).

STREAM ANALYSIS

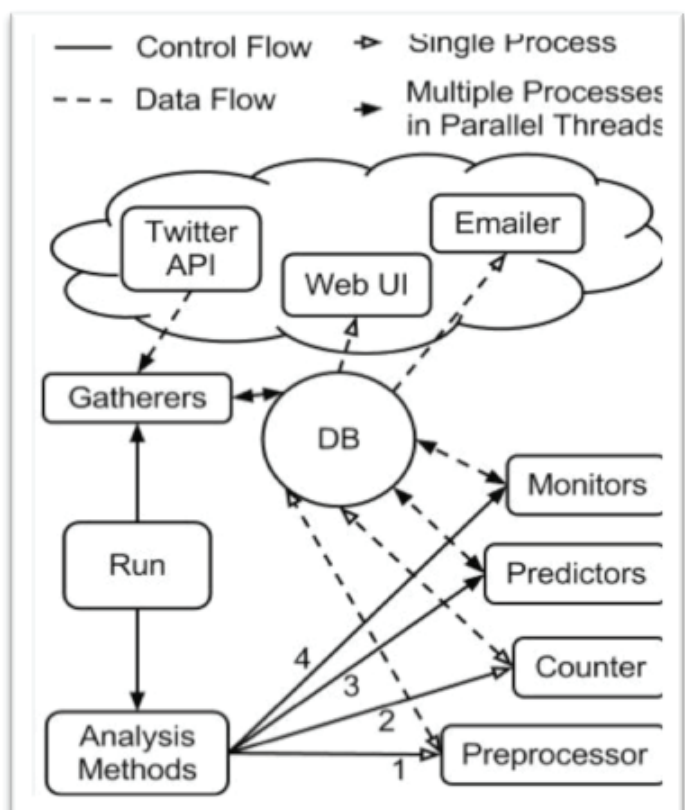


Figure 3: Netflix architecture diagram of SPOONS
Source: Augustine et al., 2012

The technique of stream analysis to conduct sentiment of a company's consumer base has become more important in our technologically driven age in which negative consumer sentiment can affect a company's brand reputation instantly. Stream analysis with Twitter has become a popular means to conduct effective customer services. One example of a company that has relied more on utilizing Twitter for stream analysis is Netflix (Figure 3). Netflix originally relied on three methods of detecting service disruptions, internal monitoring systems, external synthetic transaction monitoring systems, and general customer service operations (Augustine et al. 2012). However, Netflix customer response teams soon noticed that their current monitoring systems were not being very effective at streaming consumer sentiment. Netflix technical supporters found that manually checking their company's Twitter stream was a more effective manner to conduct customer service. Their consumers were more likely to state their service interruptions and negative experiences over Twitter than utilizing their general customer service offerings such as calling the customer service hotline. Currently, Netflix utilizes a service outage detection system called SPOONS (Swift Perceptions of Online Negative Situations) that relies on publically shared information like Twitter to conduct sentiment analysis of their consumer base and has proven to be an effective tactic to conduct customer service over their previous methods (Augustine et al. 2012).

Additionally, the utilization of stream analysis has been applied to predicting changes in the stock market, most notably the Dow Jones Industrial Average. Bollen et al., (2011) analyzed Twitter feeds through two sentiment-tracking tools, OpinionFinder and Google-Profile of Mood States. The Google-Profile of Mood States measures sentiment in six dimensional levels: calm, alert, sure, vital, kind, and happy. Bollen et al., (2011) found that the collaboration of OpinionFinder and Google-Profile of Mood States as tools to assess sentiment sources of consumers within the United States had an 86.7% ability in predicting correctly the up and down changes of the Dow Jones Industrial Average.

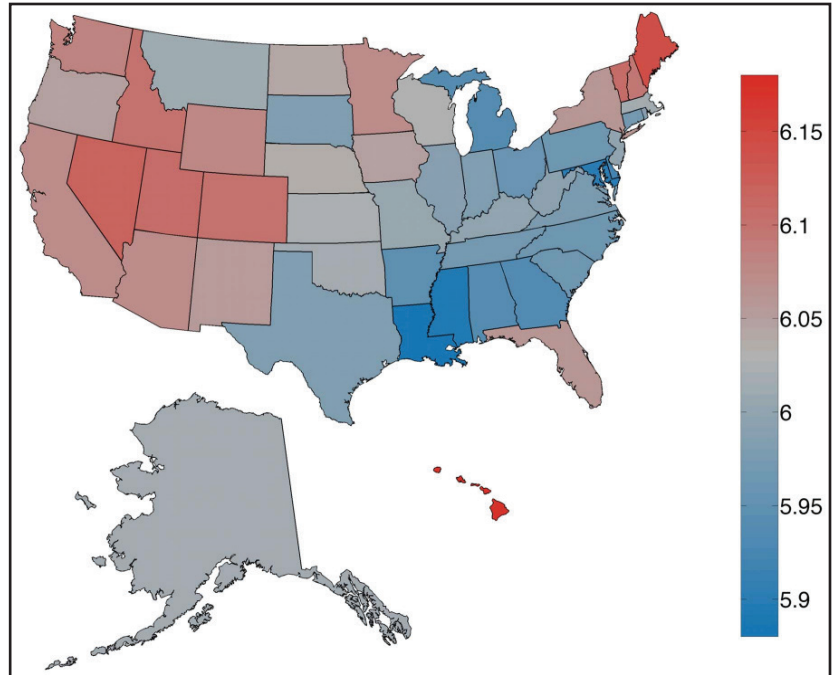


Figure 4: Average word happiness for geo-tagged tweets in the US.

Source: Mitchel et al., 2013

GEO-MAPPING SENTIMENT ANALYSIS

An emerging trend that would prove beneficial within the competitive intelligence landscape is connecting Twitter sentiment with what Mitchell et al., (2013) call the geography of happiness, an innovative method of geo-mapping Twitter sentiment data (Figure 4). The authors sought to investigate how the geographical place correlates with societal levels of happiness, most specifically in the urban centers in the United States. Mitchell et al., (2013) coupled a 80 million word data set from Twitter and annually surveyed characteristics from U.S. census data from all 50 states and 400 urban centers in order to generate an overall happiness taxonomy for United States consumers. It was found that the happiest states were: Hawaii, Maine, Nevada, Utah, and Vermont, with the saddest states being Louisiana, Mississippi, Maryland, Delaware, and Georgia. A cultivation of Twitter data streams with U.S. census data to graph geographic sentiment is an area that competitive intelligence professionals could apply all the way from strategic initiatives down to the tactical levels of planning. The type of method described by Mitchell et al., (2013) has the applicability to examine both macro and micro environmental factors with this geographical sentiment analysis technique.

PERSONAL APPLICATION OF SENTIMENT ANALYSIS WITH TWITTER

To conduct topic-based sentiment, the author utilized a paid subscription tool to convey more clearly consumer sentiment between the clothing companies of PacSun and Aeropostale, a technique of analyzing topic based consumer sentiment. Top hashtags and words were displayed in visualizations based on their frequency of occurrence. Twitter data for both companies was gathered every hour by the utilized application for a duration of two months.

Many of the top words associated with Aeropostale’s Twitter account were negative in connotation and even included a few derogatory terms (Figure 5). Going through many of the collected Aeropostale Twitter feeds confirmed that these top words were used when conveying displeasure with the clothing line’s low quality or unpopularity. Moreover, analysis of top hashtags associated with Aeropostale demonstrated similar negative sentiment (Figure 5). The hashtags identify that the Aeropostale clothing line is unpopular among Twitter followers and highly likely to be unpopular among their target consumer base of 14-17 year olds. Further analysis of the collected Aeropostale Twitter feeds demonstrated that even when hashtags #aero, or #aeropostale were mentioned in a tweet it was likely mocking the clothing line for its unpopularity, especially among their target demographic of customers.

Sentiment analysis of PacSun’s Twitter account emphasized more positive consumer engagement towards the PacSun brand (Figure 5). Through an analysis of the collected Twitter feeds these words were emphasized in tweets in which consumer’s were excited about the new clothing lines released for the summer of 2013, quality of the PacSun clothing line, the popularity of the produced clothing, and consumer’s love for the brand. A more significant emphasis of PacSun’s consumer sentiment is

conveyed in top hashtags associated with the company’s Twitter account, which further provided evidence of increased positive consumer sentiment towards PacSun’s products and services (Figure 5). Moreover, the Twitter page of @Young_PacSun and hashtag of #kandk4pacsun are associated with clothing lines that target both male and female demographics from middle school to college. Both @Young_PacSun and #kandk4pacsun showed up the most frequently in analyzed hashtags and consumers retweeting about these specific PacSun clothing lines. Thus, PacSun along with marketing to older demographics through the analysis of keyword searches has also reached the younger demographic of potential customers through their Twitter account by the increased popularity in the @Young_PacSun Twitter page and hashtag of #kandk4pacsun.

CONCLUSION/IMPLICATIONS

Overall, the utilization of Twitter to gage consumer sentiment is likely to be incorporated into business strategy in two instances. First, it will become vital for companies to utilize a customer relationship strategy (CRM), most importantly one that is more social networked based. In order to maintain a strong brand image and maintain a strong customer following, it is necessary for companies to have meaningful conversations with their customers either by answering their grievances or tailoring products and service to the customer’s conversations online. Coyle et al., (2012) found that the most efficient tool to create social customer relationship strategy would be the use of Twitter. Company’s utilization of Twitter allows them to participate directly into consumer’s conversations, whether the sentiment is positive or negative. As affirmed by Coyle et al., (2012), a social customer relationship strategy system with the analysis of Twitter was helpful in two distinct manners. Company postings on Twitter that

Sentiment	Aeropostale	PacSun
Top Keywords	Many derogatory words, didn't, stop, nasty, tired, & instead	always, amazing, cute, collection, until, Kandk4PacSun, & sneak peek
Top Hashtags	#reasonswecantbetogether, #middleschoolmemories, #yourbeingjudged, #sorrynotsorry, and #wheniwas14	#newcollection, #imabigfanof, #fashion, #gsom (golden state of mind), #obsessed, #fashion, #inlove, #shop, #shopping, & #gsomcollege

Figure 5:Aeropostale and PacSun top hashtags and keywords

helped to solve consumer problems were able to create strong perceptions of company trustworthiness. Secondly, an increased amount of company postings dealing with customer's grievances was more likely to generate higher empathetic feelings towards the company's brand image (Coyle et al., 2012).

Microblogging will increase in importance as a media platform for the consumer to speak their grievances about their company's product and services offerings. Calling or sending an email is considered too slow a method to accomplish this so consumers have turned to microblogging platforms like Twitter to post their grievances in an instant. Jansen et al., (2009) found in their analysis of 149,472 micro-blog postings that around 80% contained branding comments that were either information sharing or seeking focused. Nearly, 20% of microblog postings contained some sort of sentiment in which 50% were positive and 33% demonstrated negative sentiment towards company brands or services (Jansen et al., 2009).

The future of Twitter sentiment analysis is utilizing the geocoded properties of the customer's Twitter location to achieve location-based knowledge of an individual's daily life activities. This will be able to provide more insight than that which general census based data provides to business professionals. For example, Andrienko et al., (2013) geographically displayed space-time visual analytics utilizing geo-located Tweets within the greater Seattle area (See Figure 6). The authors discovered the thematic tweeting behavior of individuals within Seattle as it related to their regular activities and habits within repetitive spatial and temporal distribution patterns. In one such application, Andrienko et al., (2013) graphed the spatial distribution of tweets that had the word transportation as a keyword within the greater Seattle area. The tweets were displayed according to their frequency of occurrence within that area, with larger clusters representing a higher volume of tweets associated with the keyword transportation.

The applicability of analyzing geocoded tweets allows for endless applications within the business environment. The quick time microblogging characteristics of Twitter will allow a company to know instantaneously what customers are thinking about their products and services but, most importantly, where these customers are located geographically. Additionally, businesses will be able to utilize geocoded tweets combined with sentiment analysis to target a new consumer base with their products and services in the locations with the greatest frequency

of tweets under certain topic streams. Lastly, Twitter will be utilized as a tool to monitor all aspects of their competitors' activities. These activities will include monitoring what their competitor's consumers are talking about, the sentiment regarding these conversations, and where these conversations are located. In summation, all three sentiment analysis strategies with the utilization of Twitter (i.e. topic based sentiment, stream analysis, and high density geo-mapping) can provide tremendous business intelligence to a company and can help develop social media strategies that will attract consumers away from their competitor's products and services.

REFERENCES

- Agnihotri, Raj, Kothandaraman, Prabakar, Kashyap, Rajiv, & Singh, Ramendra (2012). 'Bringing "social: into sales: The impact of salespeople's social media use on service behaviors and value creation,' *Journal of Personal Selling & Sales Management*, September 23.
- Andrienko, Gennady, Andrienko, Natalia, Bosch, Harald, Ertl, Thomas, Fuchs, Georg, Jankowski, Piotr, & Thom, Dennis (2013). 'Thematic patterns in georeferenced tweets through space-time visual analytics,' *Computing in Science & Engineering*, May 31.
- Andzulis, James, Panagopoulos, Nikolaos G., & Rapp, Adam (2012). 'A review of social media and the implications for the sales process,' *Journal of Personal Selling and Sales Management*, July.
- Augustine, Eriq, Cushing, Cailin, Dekhtyar, Alex, McEntee, Kevin, Paterson, Kimberly, and Tognetti, Matt. (2012), 'Outage detection via real-time stream analysis: Leveraging the power of online complaints,' *WWW 2012 Companion*, April 16-20.
- Baker, Bill (2009). 'Your customer is talking to everyone,' *Information Management*, April 13.
- Ballve, Marcelo (2013). 'Earned media and social media: How brands can get beyond the hype,' *Business Insider*, July 11.
- Bhapkar, Neil (2013). '8 KPIs your content marketing measurement should include,' *Content Marketing Institute*, February 3.

- Bollen, Johan, Mao Huina, & Zeng, Xiaojun. (2011). 'Twitter mood predicts the stock market,' *Journal of Computational Science*, February 2.
- Bughin, Jacques, Chui, Michael, & Manyika, James (2012). 'Capturing business value with social technologies,' *McKinsey Quarterly*, November.
- Coyle, James R., Smith, Ted, & Platt, Glenn (2012). "I'm here to help" how companies' microblog responses to consumer problems influence brand perceptions,' *Journal of Research in Marketing*, January 5.
- Hao, Ming, Rohrdantz, Christian, Janetzko, Halldor, Dayal, Umeshwar, Keim, Daniel A., Haug, Lars-Erik, & Hsu, Mei-Chun (2011). 'Visual sentiment analysis on Twitter data streams,' *IEEE Symposium on Visual Analytics Science and Technology*, October 23-28.
- Howell, Katy (2010). 'Social media are you still ignoring it?' *MarketingWeek*, September.
- Jansen, Bernard J., Zhang, Mimi, Sobel, Kate, & Chowdury, Abdur (2009). 'The commercial impact of social mediating technologies: Micro-blogging as online word-of-mouth branding,' *Journal of the American Society for Information Science and Technology*, July 6.
- Java, Akshay, Song, Xiaodan, Finin, Tim, & Tseng, Belle (2007). 'Why we Twitter: Understanding microblogging usage and communities,' *SNA-KDD Workshop*, August 12.
- Kumar, Akshi, & Sebastian, Teeja Mary (2012). 'Sentiment analysis on Twitter,' *International Journal of Computer Science Issues*, July 01.
- Larson, Keri, Watson, Richard T. (2011), 'The value of social media: Toward measuring social media strategies,' Thirty Second International Conference on Information Systems, Shanghai 2011.
- Lipsman, Andrew, Mudd, Graham, Rich, Mike, & Bruich, Sean (2012). 'The power of "like" how brands reach (and influence) fans through social-media marketing,' *Journal of Advertising Research*, March.
- Mitchell, Lewis, Frank, Morgan R., Harris, Kameron Decker, Dodds, Sheridan Peter, & Danforth,
- Christopher M (2013). 'The geography of happiness: Connecting Twitter sentiment and expression, demographics, and objective characteristics of place,' *PLOS ONE*, May 29.
- Narayanan, Madhavan, Asur, Sitaram, Nair, Arun, Rao, Sachin, Kaushik, Avinash, Mehta, Dina, Athalye, Shubhangi, Malhotra, Amita, Almeida, Akhil, & Lalwani, Rajesh (2012). 'Social media and business,' *Vikalpa: The Journal for Decision Makers*, October-November 2012.
- Park, Jaram, Cha, Meeyoung, Kim, Hoh, & Jeong, Jaeseung (2012). 'Managing bad news in social media: A case study on Domino's Pizza crisis,' Association for the Advancement of Artificial Intelligence.
- Trainor, Kevin J. (2012). 'Relating social media technologies to performance: A capabilities-based perspective,' *Journal of Personal Selling & Sales Management*, September 23.
- Zhang, Lei, Ghosh, Riddhiman, Dekhil, Mohamed, Hsu, Meichun, & Liu, Bing (2011). 'Combining lexicon-based and learning-based methods for Twitter sentiment analysis,' Hewlett-Packard Development Company, L.P, June 21.

Nicholas Joseph Oram grew up in the small town of Ghent, New York. He is the son of George and Susan Oram, and the oldest of his three other siblings. Nicholas is a graduate of Le Moyne College where he received his B.A. in Criminology with a concentration in International Affairs. He is currently a second year graduate student at Mercyhurst University, where he is in pursuit of his M.S. in Applied Intelligence. Nicholas specializes in competitive intelligence and market research endeavors. He has conducted tremendous amounts of research into social media marketing and analysis practices and their applicability for both the B2B and B2C marketplaces.