UNLOCKING THE INSIGHTS: DATA-ANALYTIC EXAMINATIONS OF SOCIAL
SIGNALS IN ORGANIZATIONAL PRACTICES

BY

UTKU PAMUKSUZ

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Business Administration
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

      Associate Professor Ramanath Subramanyam, Chair
      Professor Michael J. Shaw
      Assistant Professor Eric C. Larson
      Assistant Professor Wencui Han

ABSTRACT

Using analytics to understand social data is an emerging research field across different academic disciplines including Information Systems. Information Systems (IS) discipline places the social media-related research at the intersection of people, organizations, and technology. While the recent focus of design science utilizes advanced data mining methods to develop decision-aiding frameworks and extend organizational boundaries by designing novel artifacts; the behavioral focus develops and tests theories that elucidate behaviors at the individual or organizational levels. This thesis comprises three essays aiming to contribute to IS research by both leveraging from the two foundational principles of IS -behavioral science and design science- and applying advanced analytics to further understand the executive, stakeholder, and brand relations.

The first paper in this thesis is addressing the problem of impression management in social media networks (SMN) in the top executive realm. This chapter theoretically analyzes the impression management strategies of top executives in SMN and investigates the implications for career success as an internal factor. Inductive machine learning techniques are used on SMN posts of top executives to answer the following research question: How is the usage of impression strategies in SMN contexts associated with top executives' career success? Specifically, we aim to understand the kinds of impression strategies (among ingratiation, intimidation, self-promotion, exemplification, supplication) that are particularly effective. While our findings suggest support for certain theoretically-proposed dimensions such as self-promotion and exemplification, one of our surprising findings is that ingratiation can be detrimental in such contexts.

The second paper focuses on the public reputation of executives in online platforms and investigates its implications for managerial outcomes as an external factor. Specifically, we examine how externally established executive reputation, in the form of word-of-mouth, may

affect managerial survivability. The recent advances in information technology require corporate executives to manage and monitor their personal reputation in the eyes of internal and external stakeholders on various communication platforms. We aim to develop theoretical and empirical support for the concept that external cues of executive reputation lead to consequences for a top executive's career path. In this chapter, we first analyze the credibility of word-of-mouth and fundamental differences of social media networks from traditional news media as an information source, and then provide detailed dimensions of the reputation concept formed through separate information sources from both corporate and executive perspectives. Second, we illustrate the outcomes of executive reputation formed in social media networks while focusing on the consequences of such reputation in the corporate world. Finally, we apply multiple data mining techniques to quantify the effects between executive reputation and managerial survivability.

The third paper presents a computational method for analyzing the personality of brands in social media networks. Compared with the wealth of research focused on automated human personality assessment, surprisingly little research has focused on advancing methods for obtaining brand personality from social media content. Social branding has become an essential form of marketing communication to convey core brand personality. Ability to use an effective marketing communications strategy to distinguish itself from competitors has become a requisite to enhance customer relationship and foster brand equity. However, for a firm to convey its intended brand personality to target audiences through SMN, it must have some capability to understand how to signal appropriate brand personality dimensions through their content generation and interactive dialogue with their consumers. Brand personality is a nuanced aspect of the brand that has a consistent set of traits aside from its functional benefits. In this study, we introduce a novel, automated and highly generalizable data analytics approach to extract near real-

time estimates of brand personalities in social media networks. Our new approach uses a hybrid machine learning algorithmic design, which bypasses often extensive manual coding tasks, thus providing an adaptable and scalable tool that can be used for a range of management studies.

In this dissertation, we make new contributions to two foundational principles of IS - behavioral science and design science- and apply advanced analytics to unlock the hidden relations of executives, stakeholders, and brands. These implementations allow us to elegantly capture several important features of the social signals that have been employed in online networking platforms. In the first two chapters, social signals are examined as internal and external associations of the executives' career outcomes. The third chapter investigates the social signals from brands' perspectives. While the first two chapters highlight and examine the relationships of several constructs, the main emphasis of the third chapter lies on designing a valuable information systems tool to complement the theoretical studies and enhance the practical implementations.

To JIN UK KIM…

Table of Contents

List of Tables

List of Figures

Chapter 1.

UNDERSTANDING THE ROLE OF IMPRESSIONS IN EXECUTIVE CAREER SUCCESS: A DATA-ANALYTIC EXAMINATION

ABSTRACT

Social Media Networks (SMN) can be leveraged by executives to create and sustain impressions on internal and external stakeholders, which has implications for career success. Unlike actors in traditional impression management (IM) contexts that focus on a single targeted audience, actors employing IM tactics in SMN must consider multiple audiences at the same time. While top executives might serve as firm spokespersons, they often simultaneously behave as individuals in a connected world who are concerned with their own careers. From a theoretical standpoint, we focus on the usage of IM tactics by top executives and re-examine the nature of the relationship between these tactics and career success in the context of SMN. Overall, we seek to answer the following question: How is the usage of IM tactics in SMN associated with top executives' career success? Specifically, we aim to understand the kinds of IM tactics (among ingratiation, intimidation, self-promotion, exemplification, supplication) that are particularly effective. We apply inductive machine learning techniques on publicly available SMN posts of top executives to answer this question. While our findings suggest support for certain theoretically-proposed dimensions such as self-promotion and exemplification, one of our surprising findings is that ingratiation can be detrimental in such contexts.

Keywords: social media networks, impression management, executive career success, text mining, learning algorithms

INTRODUCTION

"Successful leaders will no longer be measured just by stock price. Managing and communicating with shareholders, employees, government, community, customers, will be table stakes in the future. They are talking about your business anyway. Why not be included in the conversation?"

Peter Aceto, CEO of ING DIRECT Canada

In recent years, social media networks (SMN) have become a core part of online social interactions (Kane, Alavi, Labianca, & Borgatti, 2014). SMN are online platforms where users interact with others and create social networks. Facebook, LinkedIn, and Twitter are among the best known of these sites. Given that millions of people use these online platforms, participants in these vast networks have a potential to influence large audiences at once (Schniederjans, Cao, & Schniederjans, 2013). In addition to providing considerable networking potential for individuals, these platforms noticeably enhance communication between firms and key external stakeholders including investors and customers.

Through SMN, organizations and individuals can enhance, sustain, and defend their image (Li & Shiu, 2012; Mangold & Faulds, 2009). While organizations benefit from SMN as an efficient means to gain a broader audience (Ellison & boyd, 2013), individuals could also leverage SMN to further their own career success. Moreover, since top executives can serve this dual role of furthering the firm's goals and simultaneously enhancing their personal careers, they serve as influential individuals that can bridge distinct and important audiences. A recent survey from IBM notes that the active participation of top executives in SMN could more than triple by 2017 (IBM, 2012). Research in information systems could benefit from deeper understanding of underlying mechanisms that explain top executive influence in SMN and potential effects not only on firm practices, but also on executive career paths.

2

SMN provide at least three mechanisms for top executives to create favorable public image for themselves and the firms that they represent. First, SMN have tended to provide critical windows into what executives actually think. For instance, executives often write about their lives, convey their opinions on variety of topics, and discuss trending issues. Second, executives frequently voice their opinions about products or services they use which might advance the marketing objectives of the firms they represent. Third, executives can shape company news and engage with stakeholders through SMN by utilizing their influence. Overall, it is important to understand how executives can leverage SMN to form positive impressions in the eyes of their stakeholders.

We draw on the impression management literature to examine the executives' influence on the internal and external audience of a firm (Bolino, Kacmar, Turnley & Gilstrap, 2008). To apply an impression management (IM) framework, we consider top executives as actors whose personal and corporate roles engender wide audiences including friends, employees, shareholders, journalists, and board of directors. IM refers to efforts of an actor to generate, defend or otherwise alter an image held by a target audience (Goffman, 1959). Jones and Pittman (1982) identified five tactics of IM: ingratiation, supplication, self-promotion, exemplification, and intimidation. Ingratiation involves individuals rendering favors, conforming opinions, or using flattery to gain appreciation from the target audience (Bolino & Turnley, 1999; Wayne & Liden, 1995). Supplication refers to the tactics of an actor who highlights her/his own weakness to invoke empathy from the audience (Jones & Pittman, 1982). Self-promotion refers to the efforts of an actor who aims to be seen as competent and respected (Bolino & Turnley, 1999). Exemplification tactics are demonstrations of self-sacrifice for the company or community to portray moral worthiness (Bolino & Turnley, 1999). Finally, actors use the intimidation IM tactic to present

3

themselves as powerful and dangerous figures that are capable of harming a target (Mohamed, Gardner, & Paolillo, 1999).

Applying an IM lens to the executives' management of impressions in SMN requires us to understand the idea of audience boundaries. Actors, in traditional media settings, often use boundaries between separate audiences to customize their message (Leary & Kowalski, 1990). For example, an individual using email can segregate audiences by addressing a specific individual or group. In SMN, the boundaries bifurcating the audiences become more permeable since actors share information with multiple audiences simultaneously leading to *context collapse*. Context collapse refers to the overlap of an individual's multiple audiences into one single platform (boyd, 2008; Marwick & boyd, 2010). For instance, ingratiation towards a specific person might be intended to be private. However, if actors engage in ingratiation behavior in SMN, the intention of the actor will be visible to multiple audiences rather than only intended audience. Due to such situations, actors are likely to find it more difficult to manage separate impressions across their target audiences in SMN than in traditional media settings (Marwick & boyd, 2010).

Top executives are responsible for firm direction and are expected to participate in social platforms for furthering a firm's goals as well as for their personal careers (e.g., Chun, 2005; Davies & Miles, 1998, Diga & Kelleher, 2009). Our interest is in examining usage of IM tactics by top executives in SMN and their career success. Top executives are likely to be perceived as representatives of the firm by both external and internal stakeholders. An impression management perspective would suggest that executives may be motivated to serve two causes - the corporate goals and personal benefits (e.g., Rosenfeld, Giacalone & Riordan, 1995). While stakeholders would prefer that their top executives are effective communicators in social platforms who further corporate goals, top executives themselves might be driven by personal goals and might share their

thoughts, celebrity lifestyles, or political views. Due to these competing goals, top executives'engagement in SMN is different from typical users.

Context collapse presents a challenge for executives serving such competing goals. As a case in point, Sir Richard Branson, chairman of Virgin Group, frequently releases information about his family, trips and political interests through SMN. As of November 2015, greater than six million Twitter users follow Sir Richard Branson; the audience likely represents those interested in receiving information about the Virgin Group as well as those interested in his celebrity life style. Catering to such an unprecedented audience, including stakeholders, potential customers, and even prospective employees, requires that executives excel in managing online impressions.

In our work, we seek to answer the following research questions: does top executives' usage of IM tactics in SMN relate to their individual career success? If so, does usage of specific IM tactics associate with career success? Kane et al. (2014) highlight four important characteristics of a SMN: users (1) have a distinctive user profile; (2) have access to digital content through the platform; (3) can provide a list of other users with whom they maintain a relational connection; and (4) can view and traverse other users' connections. In this paper, we investigate Twitter as the primary platform since it possesses these four characteristics and the fact that communications are predominately public.

We intend to contribute to the IM literature by investigating IM usage in SMN and the applicability of prior established relationships. We apply contemporary data analytic methods to answer our research questions and test the robustness of the observed relationships. The remainder of this paper is structured as follows. We next summarize prior literature regarding impression management usage and separately examine research at the firm level and individual level in traditional media and in social media networks. In the following section, we present our theoretical

model and hypotheses. Next, we describe our methodology, data collection procedure and operationalization of our measures followed by our empirical results and concluding remarks.

PRIOR LITERATURE

Previous literature has extensively examined individual tactics and firm level impression management strategies in face-to-face settings and traditional media. Researchers have focused on actors engaging in IM tactics mainly in a face-to-face context and, to a lesser extent, on mediated interaction such as print, phone calls, or emails in the management of impressions (e.g., Rettie, 2009). From a firm perspective, researchers have also examined how organizations used IM strategies through annual reports, broadcast media, and special organizational programs (e.g. Elsbach, Sutton & Principe, 1998). We categorize IM research along two dimensions: type of media and unit of analysis. Figure 1.1 depicts our area of focus and a simple viewpoint of IM literature. We next discuss literature at the firm level and follow it with a discussion at the individual level.

	Traditional Media	SMN
FIRM-LEVEL IM	☑	☑
INDIVIDUAL IM (Top Executives)	☑	Our Research

Figure 1.1: Impression management usage at the individual and firm level across different media platforms.

Impression Management at the Firm Level

In the context of traditional media, impression management research has traditionally focused on how organizations employ IM tactics to enhance, maintain and defend a positive image with a variety of stakeholders outside and inside the organization (Bolino et al., 2008; Elsbach et al., 1998). For instance, it has been observed that firms frequently employ IM strategies in response to a tarnished organizational image, in order to influence external impressions and representations of the organization in annual reports and mass media (Sutton & Callahan, 1987). Conlon & Murray (1996) explored defensive and assertive IM strategies used by firms to manage external impressions.

Emergence of SMN, such as Facebook, Twitter, and LinkedIn, has created novel online platforms for firms to interact with stakeholders. It has become essential that firms understand the dynamics of how SMN can be used to create and to sustain a preferable image in the eyes of the public. For instance, Luo, Zhang, & Duan (2013) found a positive relationship between SMN activity of consumers and firm equity value. Along similar lines, Schniederjans et al. (2013) found that organizational IM strategies in social media networks have a positive association with a firm's financial performance.

Impression Management at the Individual Level

At the individual level, researchers have examined IM to understand ways in which the behaviors of actors influence audiences. Researchers have studied the role of impression management in individual relationships and activities within organizational settings, such as supervisor-subordinate relations (Fandt & Ferris, 1990; Yukl & Falbe, 1990), employee selection (Kacmar et al., 1992), and performance evaluations (Wayne & Kacmar, 1991). This literature has primarily examined face-to-face communication to draw theoretical implications.

Constructive activity on SMN can complement individuals' professional image within and outside of organizations. We expect that the primary motivation behind individuals engaging in IM tactics in online platforms is the public nature of the actors' image, which can enhance or diminish one's career success depending on the extent to which one manages online impressions carefully (e.g., Tong, Van der Heide, Langwell & Walther, 2008). As an illustration of the latter effect, recently, researchers observed that negative posts on social networking platforms were found to impact hiring decisions (Bohnert & Ross, 2010). Career related concerns might be exacerbated within a competitive environment for top executives.

Management scholars have examined how top executives influence the perception of internal and external stakeholders in mass media (e.g., Carter, 2006). In a complementary stream of work, researchers have investigated the effects of top executives' IM usage in traditional settings and found that IM tactics employed by top executives can enable career success (Westphal & Stern, 2007; Westphal & Stern, 2006). IM strategies were found to be not only employed toward board members and peer directors, but also toward a wider range of audiences including shareholders (Godfrey, Mather, & Ramsey, 2003) and journalists (Westphal & Deephouse, 2011). Since much of this body of work has been studied in the context of traditional and face-to-face settings, our goal is to re-examine some of these established relationships, given the open and public nature of SMN and the unknown effect of context collapse.

THEORY AND MODEL

Our dependent variable of interest is career success. Career success is described as the collected positive work and psychological outcomes resulting from one's work experiences (Seibert, Kraimer, & Liden, 2001). Career success has traditionally been examined as a combination of objective outcomes such as pay ascendancy and promotion and subjective elements such as job

8

and career satisfaction (Hughes, 1937). A majority of prior research on career success typically has focused on objective outcomes (e.g., Gutteridge, 1973) rather than subjective outcomes (Gattiker & Larwood, 1989). Objective career success can be operationalized through verifiable measures such as total earnings. Overall, our survey of existing research suggests that objective elements are widely considered as verifiable indicators of career success (Heslin, 2005).

Among predictors of career success, impression management has been observed to be an important driver of career success (Bolino et al., 2008; Wayne & Ferris, 1990; Wayne & Liden 1995) especially if used by top executives (e.g., Singh & Vinnicombe, 2001). Top executives straddle both organizational and individual roles which enables us to integrate insights from both organizational and individual-level literature on impression management. We adopt Jones and Pittman's (1982) IM taxonomy, validated by Bolino and Turnley (1999), which applies to both the individual and organizational level of analysis. We specifically analyze the relationships between the dimensions of IM and career success. These relationships are depicted below in Figure 1.2. We intend to revisit the role of these tactics when applied to the SMN context where IM tactics are visible not only to the target audience but also to other unintended observers. We propose that the employment of IM tactics across this collapsed audience has a non-trivial association with career success of executives.

We expect that when the definition of audience and purpose of IM are re-visited in the SMN context, some of the traditional lines of thought fall apart or need re-examination. The relationship between each of the five IM tactics and career success construct will be covered in detail as we develop the hypotheses.
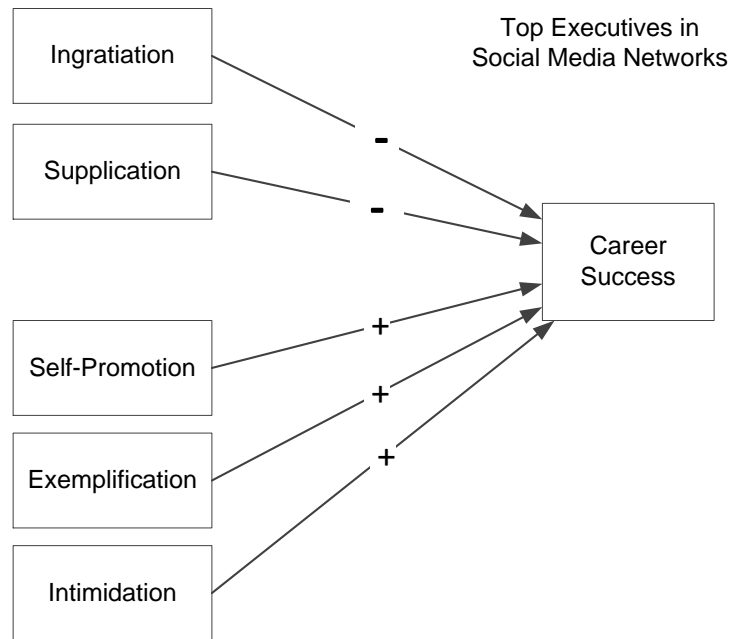
Figure 1.2: IM usage in SMN and its impact on career success

Ingratiation

Ingratiation involves individuals rendering favors, sycophancy, or using flattery to gain appreciation from the target audience (Bolino & Turnley, 1999; Wayne & Liden, 1995). Most theorists propose that ingratiation is a common and often successful method of social influence (e.g. Bolino, 1999; Gordon, 1996; Leary & Kowalski, 1990; Rosenfeld et al., 1995; Wayne & Liden 1995). By engaging in ingratiation, the employee limits the supervisor's choices to punish and control her/him (Rosenfeld et al. 1995).

Westphal and Stern (2006) suggest that top executive IM usage is a substitute for social and educational background factors that affect career success. Usage of IM tactics toward board members and peer directors has been associated with top executives receiving board appointments. Further, Westphal and Stern (2007) find that gender and ethnicity moderate the relationship between IM tactics usage and career success. In traditional media settings, these tactics were found to be employed not only toward board members and peer directors but also toward a wider range of audiences including shareholders (Godfrey et al., 2003), financial analysts (Westphal & Graebner, 2010), and journalists (Westphal & Deephouse, 2011).

Note that previous research shows that engagement in ingratiation by top executives is mostly applied toward a specific target audience such as board of directors, peers, shareholders, or journalists. However, the possibility of context collapse in SMN exists because of inherent asynchronous and many-to-many communication structure. Due to the context collapse, distinct audiences may react differently. The usage of ingratiation may be perceived as favorable by the target audience, such as a corporate board member who is motivated by vanity (Vonk, 2002), or financial analysts who are motivated by access to information (Westphal & Graebner, 2010)

However, other audience groups and bystanders may judge this behavior as dislikeable and manipulative (Vonk, 1998). When a bystander observes an exchange between an actor and a target, a bystander is likely to question the validity of ingratiatory behavior (Gordon, 1996). Such bystanders could outnumber the targets of ingratiation attempts. In addition to these bystanders, individual members within a targeted group audience might also feel resentment. Such aggregated dislike towards a top executive is likely to affect a firm's corporate image (Wade, Porac, Pollock & Graffin, 2006). Considering the substantial influence of various stakeholders on firm practices

and executives' careers (Berman, Wicks, Kotha & Jones, 1999), we propose our first hypothesis as follows:

Hypothesis 1: Top executives' engagement in ingratiation on SMN has a negative association with their career success.

Supplication

When actors engage in supplication, they highlight their own weakness to influence the audience (Jones & Pittman, 1982). By pointing to their incompetence, the actors are attempting to generate an image of being needy and dependent. The goal of supplication is to receive help to complete a task or avoid additional assignments. Supplication may decrease the perceived level of competence and is often used within organizations only as a last resort (Turnley & Bolino, 2001). However, the supplication tactic may make supervisors feel superior to others (Jones & Pittman, 1982). Longenecker, Sims, and Gioia (1987) likewise propose that the supplication tactic may make supervisors feel pity which generates sympathy toward the employee. Further, Bolino and Turnley (2003) find that seeking assistance is viewed positively in work environments.

Previous findings show that at the firm level, supplication is utilized by organizations to find solutions to cure their emerging problems (Mohamed et al., 1999). For instance, Frito-Lay launched a 'Do us a Flavor' campaign soliciting consumer ideas to refresh their brand image. In addition, supplication helps organizations remedy problems quicker because of the generation of word-of-mouth in social media networks (Schniederjans et al., 2013). For example, the U.S. Army uses the hashtag "#WeNeedYou" on SMN to recruit minority applicants.

However, in the context of top executive behaviors, we posit that supplication tactics might create an undesired image in the SMN environment. In contrast to displaying competent and powerful images which might be favored by audiences (Gaines-Ross, 2000; Lucero, Kwang &

Pang, 2009; Wayne & Ferris, 1990; Yukl & Tracey, 1992), exhibition of weakness or vulnerability of a top executive may result in a loss of confidence and a negative perception in the eyes of the stakeholders. As an example from our data set, "Sorry to bombard you with tweets. Am nervous as hell. Helps [sic]" portrays a weak image of an executive and draws instant attention of different stakeholders such as the board of directors.

The board of directors who play an important role in corporate governance and meticulously monitor the top management (Weisbach, 1988) are likely to notice managerial weaknesses through SMN. SMN platforms will likely aid in perpetuating negative impressions because of high public visibility and fast information diffusion through re-posting, sharing, and liking behaviors of the audience. As a result, we posit that career benefits, such as compensation of top executives will likely decline (Jensen & Murphy, 1990) with increased use of supplication. Our next hypothesis is as follows:

Hypothesis 2: Top executives' engagement in supplication on SMN has a negative association with their career success.

Self-Promotion

The goal of self-promotion is to be seen as competent and respected (Bolino & Turnley, 1999) distinct from ingratiation, which focuses on being liked. The actor promotes his or her general abilities like intelligence, business acumen, or specific skills (Rosenfeld et al., 1995). Employees who utilize self-promotion tactics are more likely to be perceived as productive versus employees who choose not engage in such tactics (Wayne & Ferris, 1990). Likewise, researchers have found that the use of self-promotion tactics has a positive effect on interviewee's evaluations (e.g., Gilmore & Ferris, 1989).

Among senior executives, promotion of accomplishments and demonstration of

competencies would be expected of business leaders. In addition, companies have an interest in presenting their top executives as exemplary to audiences such as employees, the press, and the financial community (Pollach & Kerbler, 2011). For example, former CEO of Chrysler, Lee Iacocca, and former CEO of General Electric, Jack Welch, have often exhibited competencies by means of traditional media channels with their book publications about leadership (Welch & Byrne, 2001). However, today SMN are excellent platforms for top executives to promote personal reputation. For instance, Tim Cook, CEO of Apple Inc. has almost two million followers on Twitter which speaks to his potential immediate audience.

Top executives' self-image, highlighted accomplishments, and reputation in the public eye can be effectively enhanced and quickly disseminated to wider audiences by means of SMN. Stakeholders among the audience will respond positively to the self-promotion of top management teams and individuals (Gaines-Ross, 2000). As a result, executives could potentially gain benefits such as compensation increases (Murphy, 1999).

Hypothesis 3: Top executives' engagement in self-promotion on SMN has a positive association with their career success.

Exemplification

Exemplification can be described as demonstration of self-sacrifice for the company or community to portray moral worthiness (Bolino and Turnley, 1999). Exemplifiers let others know that they work hard and engage in self-sacrifice embracing corporate and personal social responsibility roles. Turnley and Bolino (2001) find that exemplifiers are more likely to be perceived as dedicated and industrious by peers.

Top executives employ exemplification tactics to demonstrate their corporate and personal social responsibility roles. For example, Gregg Steinhafel, CEO of Target Inc. expressed his

opinions toward a better child education system and reducing plastic bag consumption, a message which quickly spread to a broad audience via SMN. Similarly, Tesla Motors' CEO Elon Musk displayed his vision for a greener world: "@Enric_Sala Earth is some green patches surrounded by ocean. We need to protect more of it [sic]". These behaviors are examples of exemplification tactics.

SMN can provide a platform to share information about individual social activities and promote trust from audience members through greater information-sharing capabilities (Oh, Agrawal, & Rao, 2013). Thus, we expect that exemplification tactics of top executives on SMN will enable them to reach multiple audiences with ease and in a cost-efficient manner which will improve their own career success (Jensen & Murphy, 1990).

Hypothesis 4: Top executives' engagement in exemplification on SMN has a positive association with their career success.

Intimidation

Intimidation strategy is generally employed by actors in order to present themselves as powerful and potentially competitive figures that are capable of harming a target audience (Mohamed et al., 1999). Intimidation is more likely to take place in non-voluntary relationships like the one between supervisors and employees. Within organizations, intimidation is usually a form of downward influence where individuals with higher power attempt to influence individuals with lower power in the organizational hierarchy (Rosenfeld et al., 1995). Previous findings support the notion that usage of intimidation tactics by executives within the organization may accelerate the efficiency of getting a job done and may lead to a situation where the executive is perceived as more powerful (Yukl & Tracey, 1992).

However, intimidation is not limited to organizational settings. Social networking platforms provide a unique opportunity to observe relationships between top executives and members of the external audience. Intimidation signals may be observed through Twitter feeds of top executives. For example, it has been noted that T-Mobile's CEO John Legere used Twitter frequently to threaten and intimidate the competitors of T-Mobile and gleefully posted about their various mishaps (Frank, 2014). Thus, SMN are not only likely to enhance the speed of diffusion of this impression among stakeholders, but also provide an efficient platform for signaling higher power over others.

Although an intimidation strategy may not be preferred by organizations when handling internal and external entities, executives who convey a strong posture over others may be favored in crisis situations that require immediate action (Gardner & Avolio, 1998) and instant decision making (Lucero et al., 2009). An impression management perspective would suggest that top executives engaged in such tactics may be perceived as more competitive and powerful (Rosenfeld et al., 1995). Thus, we hypothesize:

Hypothesis 5: Top executives' engagement in intimidation on SMN has a positive association with their career success.

RESEARCH SETTING AND METHODOLOGY

In this section, we address the communication platform, data collection procedures, and the description and operationalization of variables used for our analysis. We first describe the social networking setting that we use for our implementation and provide details about data resources and data collection procedures. Next, we present operationalization of the dependent and

16

independent variables. Finally, we illustrate the data mining techniques employed to reflect the independent variables.

Social Media Networks Setting

Twitter launched in 2006 as a microblogging platform hosting one of the largest online communities where the users can broadcast and consume content (Kane et al., 2014). Twitter users broadcast and consume content by posting and reading 'tweets', text-based messages of up to 140 characters. We disentangle specific influence-seeking and impression-forming behaviors of executives since the observation platform is public by default, and permits researchers to examine multi-directional interactions among actors. Specifically, we harvest Twitter data to methodically investigate the use of IM tactics of top executives in SMN. Twitter is a valid source of data since the SEC ruled that messages from authenticated social media accounts are legitimate outlets for key information and in compliance with the 2013 Regulation Fair Disclosure. Top executives now consider Twitter as an alternative medium for both announcement of corporate news and sharing of personal interests. Our goal is to analyze impression management strategies manifest in Twitter messages from executives' Twitter accounts.

Data Collection

We compile a list of all executives from Standard & Poor's (S&P) 1500 company index using the Compustat database. We use the S&P 1500 firms because of their high visibility and large investor base, which implies that this is a suitable context for examining context collapse (e.g. Hollander, Pronk & Roelofsen, 2010). Our sample contains 7,549 top executives with at least five executives per firm, the number of executives who report their compensation in the firms' annual proxy statements (SEC, 2014).

From the pool of 7549 executives, we manually inspect half of the population to identify executives with Twitter accounts. We arrive at 130 top executives who were active on Twitter as of the beginning of 2013. Next, we examine the authenticity of the Twitter accounts. Twitter verifies accounts for authenticity by posting a blue verified badge, a solid blue circle containing a white checkmark on Twitter profiles. We eliminate 19 accounts that lacked a verified badge on the executive's profile page. In addition, we eliminate one more account which was found to be a second account for an executive that was not active. In the end, we retain messages of 110 top executives' accounts for our analysis.

While an account may be verified, there is no guarantee that the source of the posts is the executive. For instance, a public relations team may use the account on behalf of a senior executive. However, this does not affect or alter our analysis direction since we are interested in outcomes of impression tactics in the eye of stakeholders in the audience. Stakeholders who see the executive's name on the account attribute the message to the executive.

Next, we assign two machines to collect streaming data by querying the posts of top executives through Twitter application programming interface (API). We run Python programing language scripts in parallel to gather all the tweets from 110 executives' accounts throughout the year 2013 until the beginning of 2014. Our dataset comprises more than 230,000 messages sent by top executives including metadata such as user-id, time-stamp and content type. The dataset contains tweets, retweets and reply-messages from executives to other Twitter users. We remove non-text tweets. Our data set contained 171,893 tweets.

We collect executive career success data from the SEC filings, SEC Form DEF 14-A, of publicly traded companies and *Standard & Poor's* ExecuComp database. Executive security holdings and their compensation packages are included in the SEC forms DEF-14A, filed annually

in the end of the fiscal year by companies following Section 14(a) of the securities Exchange Act of 1934. ExecuComp database uses SEC forms to combine salary, bonus, total value of restricted stock granted, net value of stock options exercised, and long term incentive payouts. We augment ExecuComp data set with portions of SEC filings of firms.

Dependent Variable

In our study, we are interested in yearly changes in executives' earnings associated with their careers. Previous literature explains career success of top executives along two dimensions: 1) objective elements, such as payment escalation and income increase, and 2) subjective elements, such as job and career satisfaction (Heslin, 2005). We use the objective metrics of career success because they are directly observable, measurable, and verifiable by an independent third party (Nicholson, 2000). Within this perspective, O'Reilly, Doerr, Caldwell, and Chatman (2014) used objective elements to operationalize career success which compose of total compensation including salary, bonus, annual awards, total value of restricted stock units granted, total value of stock options granted, long-term incentive payouts, and all other remuneration. In this paper, we derive the total compensation value to operationalize executive career success from ExecuComp database as the sum of all the earning elements throughout the fiscal year.

Since we focus on the effect of online IM tactics that top executives engage in over the course of the year, we investigate the change in pay from the beginning and end 2013 fiscal year which ends in 2014. We utilize percent change in earnings as our measure instead of natural logarithm of actual dollar values in order to eliminate the effect of outliers due to size differences among firms in our data set (Greene, 2003). The larger the firm measured by total assets and sales, the greater is the total pay provided to the top executives (Staw & Epstein, 2000). Thus, we

operationalize career success in our empirical model as the percentage change in total earnings from the 2012 fiscal year to the 2013 fiscal year.

<div align="center">**Independent Variables**</div>

We operationalize five dimensions of IM tactics —ingratiation, supplication, self-promotion, exemplification, and intimidation— by drawing on Bolino and Turnley's (1999) approach to assessing IM tactics. In prior work, Bolino & Turnley and others (see Bolino et al., 2008 for a review) have employed self-reported surveys to measure IM tactics. More recently, Benthaus (2014) and Schniederjans et al. (2013) proposed a framework to measure organizational IM tactics from online streaming data. Benthaus (2014) used sentiment analysis and manual content coding to extract impression strategies of financial firms. We extend Benthaus's method by using a comprehensive text mining technique that scales to larger data repositories. In our study, we employ Perceptron, a machine learning algorithm, to extract IM tactics of top executives from Twitter (Ng, Goh, & Low, 1997; Schutze, Hull, & Pedersen, 1995). See supporting information section for our rationale in choosing this algorithm.

We classify and quantify five IM tactics contained in tweets posted by top executives as our independent variables. Manually coding all tweets with IM tactics is difficult as the cost will be prohibitive and human coders may introduce bias while working with high volumes of data. Thus, we leverage semi-supervised algorithmic procedures to automatically classify IM tactics from an unstructured corpus of tweets posted by top executives. Further details about the algorithmic classification procedures and supervised learning models are provided in supporting information section. An overview of the steps to derive independent variables is summarized as in the following:

    1. Selection of training set: tweets for manual coding

2. Train human coders

3. Employ automated text classification algorithm

4. Predict and quantify IM tactics

Selection of Training Set

The first step in deriving our independent variables is to code a subset of tweets with corresponding IM tactics. In order to select a balanced and unbiased subset of tweets in a cost-efficient manner for manual coding, we follow sampling by clustering method, which uses k-means clustering to arrive at a representative sample of tweets (Zhu, Wang, Yao, & Tsou, 2008). The details of sampling-by-clustering method are provided in supporting information section.

Training Human Coders

The second step was development of a training document for human coders and a coding scheme to classify tweets into individual IM tactics. Morris (1994) tested the validity and reliability of manual coding approaches and achieved an acceptable level of semantic validity. We follow her structural procedure to classify the content based on a coding scheme and to make the results replicable by others. We define single tweets as the unit for analysis because they can be objectively recognized by the coders without losing contextual information (Harwood & Gary, 2003). We create a training document, summarized in Table 1.1 and explained further in supporting information section, which highlight behaviors, definition, and examples of how Tweets reflect specific IM tactics (Bolino & Turnley, 1999; Jones & Pittman, 1982; Mohamed, et al.1999).

Table 1.1: Impression Management Tactics Training Document

Behavior	Definition	Examples
Ingratiation	Behaviors used by top executives to make them look more favorable to an audience.	Opinion conformity, favor rendering and flattery towards stakeholders. @xxx: I'm with Arie. excited to hang out with you tomorrow:)
Self-Promotion	Behaviors presenting the top executives as highly productive, successful, and competent.	Personal shares via Twitter to promote one's competency. @xxx: If you need to get better organized and build systems, do what I do
Exemplification	Behaviors used by top executives to demonstrate their integrity, social responsibility and moral worthiness	Comments in Twitter to promote socially responsible activities. @xxx: The best thing you can learn is to forget yourself and serve the community
Intimidation	Behaviors displaying the powerful and intimidating side of personality to establish control on audience.	Personal posts of top executives suggesting their power over similar competitors or employers. @xxx: You service is absolutely terrible - get on the ball! @yyy
Supplication	Behaviors displaying an image of dependency and weakness to solicit help from others	Using Social Media Networks to gather public support and solution for problems @xxx: Ive tried for 13 yrs to fix the officiating in this league and I have failed miserably. Any Suggestions? I need help

Table 1.2: Impression Management Coding Sample

IM Tactic	Date	Content
Self-Promotion	23-Aug	Honored and humbled to be recognized today by the AAF Mosaic Awards @aafmosaiccenter.
Ingratiation	12-Nov	A big thanks to @DavidKirkpatric for bringing together great minds to consider the interplay of technology and mankind! Very good stuff!

Table 1.2: Impression Management Coding Sample (Cont.)		
Intimidation	16-Aug	If all anyone remembers are my failures, then they are *no one* to me.
Supplication	24-Jul	Please take a moment to download my new @altimetergroup report on #DigitalTransformation w/@starbucks @sephora @ford
Exemplification	19-Aug	Bring hope to the #ChildrenofSyria. Join WU + @UNICEF. Donations doubled thru @WesternUnion up to $100K total.
Null	11-Sep	Something...different…

Next, two research assistants from the authors' institution were initially trained based on the theoretical foundations (Jones & Pittman, 1982) and the taxonomy of impression management tactics (Bolino & Turnley, 1999; Mohamed et al, 1999). The research assistants code the tweets for each of the five IM tactics and a null category. Several iterative practice sessions were conducted with Twitter data sub-samples to train the coders with the content. This sub-sample of 760 tweets was only used for training of human coders and eventually excluded from the final dataset. We observed an inter-coder reliability score of 0.82 which is greater than the threshold recommended by Krippendorff (2012). We completed the initial IM strategy identification phase by manually training 3240 messages with corresponding IM tactics. A snapshot of the outcome at this phase is presented in Table 1.2. This set of 3240 messages will serve as training inputs into the learning algorithm.

Text Classification Algorithm

The third step in deriving our independent variables is the application of automated text classification methods based on previously identified IM tactics. Text classification through supervised learning techniques has increasingly been employed in mainstream information

systems literature. Recently, Abbasi et al., (2010) propose a comprehensive analysis of Statistical Learning Theory based text classification techniques to reliably detect fake websites. These techniques have been frequently applied in other contexts including detection of financial fraud (Abbasi, Albrecht, Vance, & Hansen, 2012), prediction of strategic gaming by agents (Boylu, Aytug, & Koehler, 2010), and estimation of the global status of web pages (Pant & Srinivasan, 2013). The goal of this step is to select the best classification method for our analysis from the alternatives, keeping in mind that our main priority is to minimize classification error and that our context is one where there are multiple classes for prediction.

We examined four classification approaches and present the best performing algorithm in each type of method. The classification methods could be broadly categorized as frequency-based (e.g. Naïve Bayes), proximity-based classifiers (e.g. K-Nearest Neighbor), non-probabilistic linear classifier (e.g. Support Vector Machines), and neural network based (e.g. Perceptron).

These classification algorithms are prone to errors due to high lexical and structural syntactic ambiguity, such as a word being classifiable as either a noun or verb. For example, we provided a few instances of misclassification from our data while applying learning algorithms looking at terms individually to extract features, which are measurable and identifying characteristics of each IM tactic. In the illustration below (Table 1.3), the word 'thanks' is used sarcastically in the third tweet which leads to false positive classification.

Table 1.3: Example of Misclassification

Tweet	IM Tactic	Prediction
"A thank you and congratulations to @dickc, @mgupta, @vijaya, @gabrielstricker, and the @Twitter team! $TWTR"	Ingratiation	True
A big thanks to @DavidKirkpatric for bringing together great minds to consider the interplay of technology and mankind! Very good stuff!"	Ingratiation	True
"I am late, thanks to Manhattan traffic. pic.twitter.com/pLkPLsE"	Ingratiation	False

To minimize misclassification errors, we employ a combination of classical and customized tokenization and feature extraction techniques. See details in supporting information section. Tokenization is the process of taking text and splitting it into individual terms. Feature extraction takes these set of terms and transforms those sets into numerical feature vectors. Specifically, we combine some terms as bi-gram and tri-gram word lists into similar phrases to achieve better feature vectors for IM tactic representation (Feldman & Sanger, 2007) and observed a high accuracy rate of 70.9% with Perceptron learning (see Table 4). Our validation process involves 10-fold cross validation, which provides higher rates of accuracy when the model is applied to predict from an independent dataset of tweets (Abbasi and Chen, 2008). The details of the predictive model are presented next.

Table 1.4: Comparison of Text Classification Algorithms

Classification Method	Accuracy
Perceptron (number of iterations=50)	0.709
Linear-SVM	0.664
K-Nearest Neighbors Classifier (number of neighbors=15)	0.636
Naïve Bayes	0.611

Prediction and Quantification of IM Tactics

The final step is the prediction and quantification of IM tactics from the entire data set so that we could eventually analyze the effects of IM tactics on career success. Overall flow of the training and prediction phases can be seen in Figure 1.3.

First, we clean up the raw data set by applying pre-processing to remove stop words, stemming, and punctuation and transform it to a computational format by using scikit-learn machine learning package for the Python programming language (Han, Kamber, & Pei, 2011;

Pedregosa et al., 2011). In the training phase, following the steps highlighted above section, we create a collection of labeled tweets.
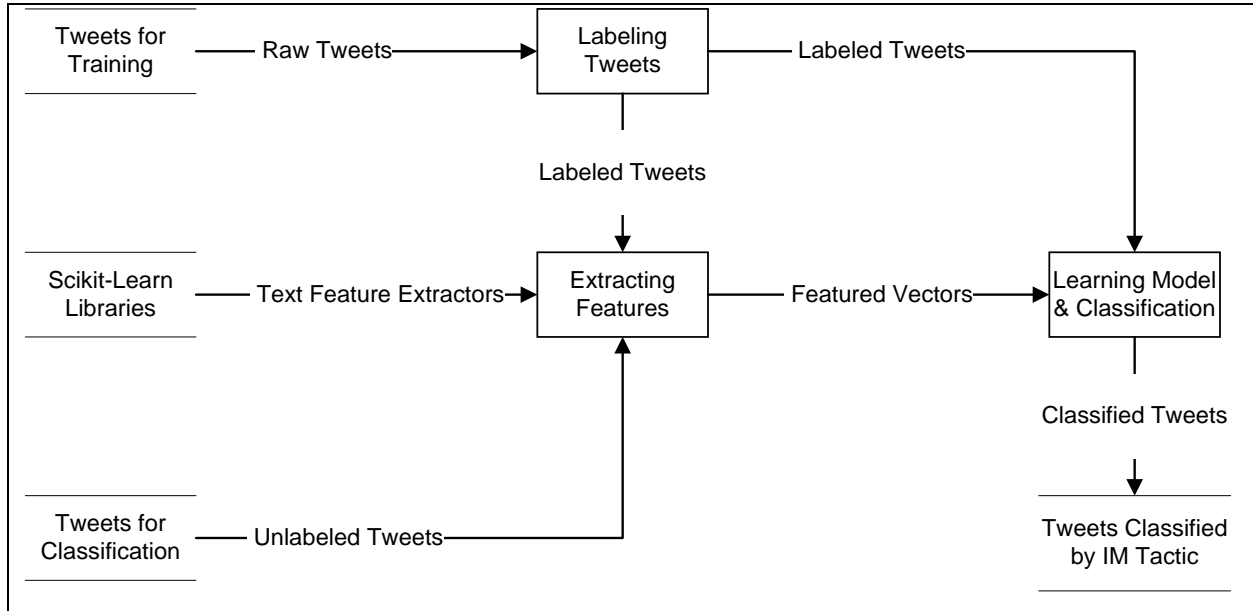
Figure 1.3: Steps involved in our classification method for IM Tactics

After the labeling process, we derive feature vectors and training set of labeled tweets consists of 3240 sample messages with 10-fold cross validation which basically partitions a given sample into 10 sub-samples and validates the accuracy of 9 subsamples over a single retained test sub-sample throughout random iterations. We evaluate the accuracy of our perceptron model as follows:

$$Accuracy = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Prediction}}$$

We observe a 70.9% accuracy rate; the percentages of usage of each IM tactic are presented in Figure 1.4. If a tweet contains more than one IM tactic, Perceptron selects the dominant IM tactic. In prediction phase, we classify the entire data set based on previously trained algorithm. Next, we estimate the impact of the five dimensions of IM tactics ingratiation, intimidation,

exemplification, supplication, and self-promotion – on career success by using top executives as unit of analysis in a multiple regression.
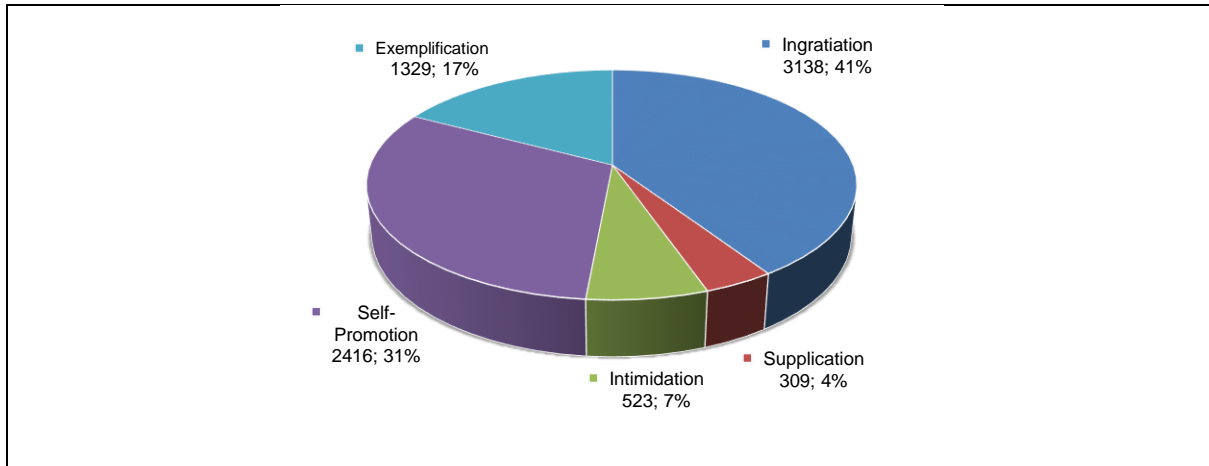
Figure 1.4: Aggregated impression management tactics and the ratio of IM Tweets

Control Variables

We control for several variables, aside from IM tactics, that could impact the executive career benefits. First, we account for industry-specific factors by using SIC codes from Compustat database to create categorical industry indicator variables in order to account for different payment patterns for top executives across industries.

Second, prior research of top executive compensation predicts a positive relationship between executive pay and corporate financial performance (e.g., Murphy, 1985). We account for non-equity but revenue-growth-based measures of financial performance. Specifically, we use sales growth and operating income growth to control for financial performance (McGuire, Sundgren, & Schneeweis, 1988). This measure is generated from Compustat database which reflects total revenue growth that has occurred from 2012 to 2013. We also control for firm size,

since larger firms are expected to pay more to their executives (Baker and Hall, 2004). Since the financial performance measures accounts for sales dimension, we employ the number of employees in fiscal year 2013 to operationalize firm size (O'Reilly et al, 2014).

Finally, we use tenure and demographic variables, such as age and gender, to account for individual differences (e.g. Lee & James, 2007). Previous research shows evidence that longer tenure in the top executive roles receive larger compensation packages than those who have shorter tenure in the executive role (O'Reilly et al., 2014). In addition, we also consider the fact that executives who founded their firms might have different rights in their firms than non-founders. We obtain this information from SEC filings of the firm and use this information in our analyses as a categorical variable.

EMPIRICAL ANALYSIS

Our empirical approach is to employ multiple linear regression analysis to estimate the effects of the five IM tactics; self-promotion, ingratiation, intimidation, exemplification, and supplication, on Career Success, accounting for the other explanatory variables. Descriptive statistics for our measures are provided in Table 1.5 below. To examine whether IM tactics are likely to cause collinearity concerns, Spearman rank correlations were computed for these measures. These correlations are shown in Table 1.6. All correlations are less than 0.5, which indicates that multicollinearity across the IM tactics is less likely (Kishore, Agrawal, & Rao, 2004).

Table 1.5: Descriptive Statistics (n=110)

	mean	sd	min	max
Career Success (%)	.677	2.12	-4.36	10.11
Ingratiation	28.52	24.42	1	98
Intimidation	4.74	5.19	1	22
Self-promotion	21.95	17.40	1	75
Exemplification	12.91	10.73	1	42
Supplication	2.80	2.68	1	18
Fin. Performance (%)	0.33	1.03	-1.03	8.29
Age	50.23	11.06	31	83
Tenure	10.10	7.32	1	35
Firm Size	43855.79	74876.83	512	317500
Founder	.51	.5	0	1
Gender	.14	.35	0	1
Industry	2.79	1.39	1	5

Table 1.6: Spearman's Correlations for IM Tactics

Ingratiation	1.00				
Intimidation	0.22	1.00			
Self-Promotion	0.35	-0.03	1.00		
Exemplification	0.20	0.08	0.39	1.00	
Supplication	-0.11	0.28	-0.05	-0.21	1.00

As a second check for multi-collinearity problems, we calculate variance inflation factors (VIF) for each variable. The average VIF values for each variable are 1.26, 1.24, 1.23, 1.22, 1.12 for ingratiation, intimidation, self-promotion, exemplification, and supplication, respectively, which are less than the acceptability threshold of 5 (Hair et al., 2006), implying that multicollinearity should not affect our results. Further, to limit potential concerns about unequal variances of our IM tactics and controls across the range of the career success measure, we employ

heteroskedasticity-consistent standard errors (White, 1980), with industry-level clustering (Rogers, 1994).

Our choice of a lagged-model, wherein IM tactics are measured in a time period prior to the time period for the dependent variable measurement, should limit concerns of endogeneity. As an additional precaution, we tested audience reach as an instrument in our model. We observed that OLS and IV models produced equivalent results (See Table 8).

Results of Multiple Regression Estimations

Using multiple linear regression analysis, we estimate the effects on career success of five IM tactics; self-promotion, ingratiation, intimidation, exemplification, and supplication. Results of the regression analysis are presented in Table 1.7. We find that the overall model is significant with an F-value of 24.6 significant at $p < 0.001$. The proportion of the variance accounted for by the model is 48.48%. We introduce variables in a step-wise manner into the multiple regression estimation. Model (1) only includes the IM tactics. Ingratiation, self-promotion, and exemplification have a statistically significant effect on top executive career success. For instance, a .0313% decrease in the managerial earnings is associated with an engagement in an additional ingratiation tactic on Twitter with $p < 0.001$. Similarly, a .0728% increase in managerial earnings is associated with an additional self-promotion tweet from the top executives with $p < 0.001$, holding other variables constant. Exemplification was also found to be significant with a .05% increase in managerial earnings. Although the signs of intimidation and supplication tactics are as hypothesized, they are not statistically significant.

Model (2) introduces control variables. We see a statistically significant relation between managerial earnings and firm financial performance which is consistent with previous findings

30

(Murphy, 1985). All other controls hold negative signs except age and they are not found to be statistically significant. Overall, the statistical significance of our explanatory variables remains after inclusion of control variables. In addition to account for industry specific effects from our model, we apply robust standard errors with industry level clustering (Rogers, 1994). As seen from the estimates presented in Model (3), we observe minimal changes in values of explanatory variables when compared to Model (2).

Table 1.7: Multiple Regression Estimates of Career Success (n=110)

	(1) Career_Success	(2) Career_Success	(3) Career_Success
Ingratiation	-0.031***	-0.026**	-0.026**
	(0.006)	(0.007)	(0.007)
Supplication	-0.048	-0.098	-0.099
	(0.058)	(0.057)	(0.073)
Self-promotion	0.072***	0.067***	0.066***
	(0.009)	(0.009)	(0.018)
Exemplification	0.049**	0.04*	0.04*
	(0.016)	(0.016)	(0.015)
Intimidation	0.01	0.039	0.039
	(0.031)	(0.034)	(0.028)
Fin. performance		0.537**	0.538**
		(0.154)	(0.184)
Age		-0.007	-0.007
		(0.017)	(0.017)
Tenure		0.004	0.004
		(0.027)	(0.027)
Founder		-0.502	-0.503
		(0.404)	(0.405)
Firm size		-0.001	-0.001
		(0.001)	(0.001)
Gender		-0.213	-0.214
		(0.496)	(0.496)
Intercept	-0.537	0.093	0.094
	(0.354)	(0.809)	(0.809)

Standard errors in parentheses $^*\,p < 0.05,\ ^{**}\,p < 0.01,\ ^{***}\,p < 0.001$

Robustness Checks and Test of Alternative Explanations

One possible counter-explanation for one of our results is that the decision of using a specific self-promotion tactic on Twitter may depend on career achievement. In essence, this would imply that one of the drivers of the self-promotion strategy engagement may be career success itself. To account for this explanation, we attempt to examine variation in self-promotion usage that stemmed from a factor unrelated to the expected effects of this strategy on career success. Specifically, we use the audience reach on Twitter as an instrument for self-promotion strategy usage decision. We combine three parameters obtained from the users' Twitter account to measure audience reach by specifically the number of the followers, the number of Retweets, and the number of mentions (Cha et al., 2013). The number of the followers directly indicates the size of the audience for a specific user. The number of Retweets implies the potential of that user to generate a self-promotion tactic with a pass-along value and the number of mentions containing the user's name indicates the capability of the user to engage other users in a conversation. We provide an audience reach score for each of the user in our data set by combining follower, retweet and mention counts into one aggregated parameter.

First of all, the most direct explanation for the relationship between self-promotion and the audience reach can be seen in the literature where self-promotion applied in traditional media channels. Basically, the actors will likely exhibit strong preferences to engage in a self-promotion activity in the arenas where they are popular. In fact, we see a strong correlation between self-promotion and audience size in our data sample (.52). However, we see little reason to expect that audience size in Twitter would affect career success of a top executive in a competitive corporate

world, other than through the efficient use of a promotion strategy to impress a target audience(1).

Above mentioned arguments need to consider two additional important facets in the statistical validation process (e.g. Murray, 2006). First of all, although a full test of exogeneity is impossible, we include audience size into our regression model. Model (4) at Table 1.8 reports a reduced form model of career success including both the instrument (audience size) and the instrumented variable (Self-promotion) as covariates. Though not a formal test of exogeneity, our results suggest that audience size has no direct effect on career success of top executives, controlling for self-promotion strategy which also confirms that the exclusion of audience size from our model.

The second concern in IV regression is to check whether the instrument is weak. The F-test for the omitted instrument is 19.2 which is sufficiently above the critical threshold and suggest that two-stage-least-squares estimates would have less than 10% of the bias observed in the OLS estimates (Stock and Yogo, 2005). Left column model (5) at Table 1.8 reports the first stage estimates for the IV models. The estimate of audience reach on self-promotion strategy is statistically significant. This result implies that audience reach strongly predict self-promotion engagement behavior of the top executives. In the second stage, the right column of the model (5) at Table 1.8 reports that the effect of self-promotion remains significant and positive in the model. Thus, to our knowledge, we can conclude that OLS and IV produce equivalent results in our dataset.

---

1 Though one might still worry that top executives leverage from their online popularity that brings career success to them, less than 5% CEO's of the Fortune top 50 companies received an income escalation who had a considerable amount of audience on Twitter in 2013.

Table 1.8 Instrumental Variable Estimates (n=110)

	(4) Reg Career_Success	(5) IV Estimation Self-promotion	Career_Success
Self-promotion	0.0640*** (3.77)		0.0596* (2.50)
Audience	0.00000243 (1.02)	-0.000116*** (-3.96)	
Ingratiation	-0.0243** (-3.08)	0.180* (2.21)	-0.0281*** (-3.54)
Supplication	-0.100 (-1.45)	0.0865 (0.18)	-0.0573 (-0.73)
Intimidation	0.0341 (1.15)	-0.573* (-2.00)	0.0262 (0.88)
Exemplification	0.0474** (3.02)	0.436* (2.40)	0.0547** (2.82)
Age	-0.00278 (-0.16)	-0.0159 (-0.09)	0.00398 (0.24)
Tenure	-0.0104 (-0.38)	-0.387 (-1.37)	-0.0310 (-1.30)
Founder	-0.223 (-0.54)	7.526 (1.97)	-0.243 (-0.73)
Gender	-0.304 (-0.60)	1.994 (0.39)	-0.116 (-0.21)

t statistics in parentheses * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

DISCUSSION OF OUR FINDINGS

For top executives, it is crucial to simultaneously manage impressions from the perspective of all the stakeholders including investors, shareholders, board of directors, peer directors, customers, employees and even prospective employees. Previous research, in traditional media settings, found that top executives indeed engage in IM tactics towards stakeholders so as to improve their personal welfare (e.g. Westphal & Deephouse, 2011; Westphal & Stern, 2006 and 2007). In the

presence of context collapse, where boundaries separating the audiences become blur and information is shared with multiple audiences simultaneously (boyd, 2008; Marwick & boyd, 2010), we observe that the way in which executives' behavior creates and sustains impressions in the eyes of internal and external stakeholders is vastly different. We find these behaviors are indeed associated in a different way with the executives' career success. A summary of our results from hypotheses testing are provided in Table 9.

Table 1.9: Summary of Hypotheses Tests

IM Tactic	Predicted Effect	Result of Hypothesis Testing
Ingratiation	Negative	Supported
Supplication	Negative	Unsupported
Self-Promotion	Positive	Supported
Exemplification	Positive	Supported
Intimidation	Positive	Unsupported

Our results summarized in Table 7 and Table 9 reveals the following. We find that only self-promotion and exemplification are positively associated with executive success. As hypothesized, we also find support for the negative effect of ingratiation. Engaging in this tactic is ill-advised and has association with career success. On the other hand, the effects of intimidation and supplication on career success are not statistically significant. Our results are in contrast to prior literature which found, in traditional media settings, that all five IM tactics were found to benefit the actor.

Ingratiation strategies towards different stakeholders at different times were found to gain career benefits (Westphal & Deephouse, 2011; Westphal & Stern, 2007). In contrast, in our setting, we find that ingratiation was found to have a negative effect on career success. Our view is that public flattery, conforming to prevalent opinion, or offering favors to a targeted audience, can all

be positively managed in traditional settings. But, the very same behaviors are likely to be perceived negatively when audience boundaries become permeable, as in SMN.

Next, our results indicate significant support for the positive effect of self-promotion on personal career benefits. Social networking platforms, such as Twitter, can facilitate portrayal of a desirable image in a fast and efficient manner. Our viewpoint is that accomplishments of executives can reach audience and cascade through stakeholder and related-influential networks in an unprecedented manner. Hence, such behavior is likely to be associated positively with personal career success.

Similar to self-promotion, the impact of exemplification, where people seek to be viewed as dedicated, and as those who go beyond the call of duty, can be considerably magnified in SMN. With exemplification, there is an added effect due to the improved potential public image rendered to the influencers (users who retweet, share and propagate such messages) due to sharing information through their networks, even if the news concerns others' good deeds.

Finally, we note that we did not find support for two hypothesized relationships. Although Schniederjans et al. (2013) find supplication strategy in social media might help firms improve financial performance, we do not find any evidence of top executive engagement in this tactic being associated with career success. One explanation is that this tactic is employed approximately 4 percent of the time. So executives might avoid this tactic fearing harm of their public image. Another explanation is that an executive who seeks help publicly from peers might seem weak, while at the same time, the benefit of being perceived as one who does not hesitate to ask for help.

Next, we find that the association between intimidation and career success is not statistically significant. Similar to supplication, this strategy is used approximately seven percent of the time. A second explanation is that the positive effect of imposing one's superiority over

36

less-confident peers is countered by the negative effect of public scrutiny over such actions, which might be perceived as a form of bullying. Third, from a firm's perspective, while it may sometimes be seen as favorable to disseminate signals of competitive advantage through Twitter, the negative effects from other stakeholders such as partner firms, rival networks, and potential customers might counter the favorable aspects of intimidation.

Theoretical Contributions

We extend traditional IM theory to the online SMNs and re-visit established relationships in this context where it is difficult to separate the audience boundaries. In this context, we find ingratiation and supplication do not create a favorable image on targeted audience. The fundamental reason is the permeable boundaries among audiences. The permeable boundaries are enabled by the social media platforms that have provided application program interfaces (APIs) to allow for integration with other websites and hardware platforms such as mobile devices (Kane et al, 2014). Because of the context collapse, the overlap of an individual's multiple audiences into one single platform, top executives today need to be more mindful about managing disparate impressions for their respective audiences in social networking platforms than in offline settings.

In addition, we seek to highlight an important concern for researchers in the area of impression management. We suggest that the outcomes of the engaged IM strategies differ substantially based on the user characteristics and thus, we caution that our results are limited to SMN usage by top executives. For example, we choose to investigate individual IM tactics usage at top executive level. While firms might be willing to see their top executives as effective storytellers in social media networks to further corporate goals, top executives themselves might find it interesting to share their thoughts, celebrity lifestyles, political views, etc. for their personal gains. Due to these competing goals, top executives' engagement in social media networks is

different from typical users who only represent themselves. Our work highlights specific and robust results in a realistic setting, in the presence of context collapse, to complement prior work on IM tactics usage in social media (e.g., Schniederjans et al., 2013) and to explicitly examine the effect of IM tactics usage of top executives on their personal career benefits (e.g., Westphal & Stern, 2006; Westphal & Stern, 2007).

Finally, we propose that self-promotion and exemplification are highly effective tactics considering the challenge of context collapse. Since these tactics do not specifically need audience segmentation for better performance, more interestingly, the complex structure of social media networks, specifically such as that of Twitter, provides a unique platform for actors to gain extended benefits from these specific IM strategies.

Practitioner Contributions

Shareholders, employees, customers and community may want their executives to be effectively utilizing social networking environments to further their causes. The potential of narrating company news, the power of utilizing a high degree of influence over discussions adjacent to their business, and engagement with internal and external stakeholders forces top executives to be adroit in communicating on SMN. While we witness an increase in the rate of SMN usage among executives from year 2012 to 2013, most of the Fortune 500 executives still hesitate to adopt such platforms and find it risky and uncomfortable (Weber Shandwick, 2014).

The emergence of numerous communication channels has created several risks for modern executives. However, companies expect their top executives to manage these risks. Our research study attempts to provide evidence about the risks and rewards of interacting with the market through SMN.

CONCLUSION AND FUTURE WORK

In this study, we specifically study public and personal types of communication preferences of top executives. Due to their leadership roles within organizations, we witness increasing levels of senior management engagement with technology through variety of professional channels, such as teleseminars, conference calls and video conferences recently. Usage of these professional channels is regulated by SEC in order to ensure disclosed information reaches to all stakeholders at the same time. Several studies have been conducted to understand the effects of such professional communications in the accounting and corporate finance literature. However, approaching this phenomenon from a different perspective, we focus on the effect of personal benefits enabled by recent advances in social information technology platforms.

As the main limitation of this study, we acknowledge that there are many other internal and external reasons that may affect managerial decision making. First, we analyze the broadcasted messages of executives and we need to ensure that whether those messages are received and regularly monitored by decision makers including board of directors. Current methodological setting do not allow us to track the recipients of broadcasted messages on Twitter. Thus, instead of explaining a casual effect between IM tactics and career success, we focus on highlighting an association between these constructs. Second, the path between IM tactic usage on SMN and career success can contain several other omitted external and internal factors including analyst assessments, scandals, stakeholder perceptions, or executive reputation in general. In this study, we focus on highlighting a potential internal effect, impression tactic usage, enabled by recent advances in social information technology platforms. These additional factors may be included in future research models to complement our work.

Another limitation of our study is the size of the data set. Although we were able to analyze about 200,000 posts of S&P 1500 executives, only ten percent of Fortune 500 CEO's are active in Twitter as of January, 2014 (Weber Shandwick, 2014). Second, we limited ourselves and made an effort to analyze five dimensions of IM on this paper to contrast with traditional settings; future research may reveal other dimensions of IM only found on SMN. Finally, our findings are bounded by our data mining approach. For instance, we rely on the given Python Scikit Learn libraries for our analysis. Note that the Python Scikit Learn libraries have been extensively used in academic research (e.g., Pedregosa et al, 2011). New approaches may prove to be more sensitive and accurate.

However, we believe our methodology is novel and timely because of the following reasons. First, we complement prior research in this stream that uses self-reported survey data. Our view is that this approach might be subject to data limitations pertaining to sample size, recall biases, and low response rates (Bolino et al., 2008). Moreover, since impression management strategies can be used unintentionally (Liden & Mitchell, 1988), it is possible to capture sub-conscious tactics by analyzing instant social network messages.

Although IM theory implies that individuals should act differently when facing multiple audiences, little empirical research has explicitly tested this hypothesis. SMN provides members a connected platform to build and sustain various social connections (Parks, 2010; Marwick & boyd, 2010), which can serve as a setting for comparing and contrasting how individuals monitor and adjust their virtual identities in the simultaneous presence of different audiences (Carter & Grover, 2015).

Finally, no study to date and to the best of our knowledge, has investigated which IM tactics used by top executives in online settings. Interestingly, there are certain anecdotal stories

and even handbooks written on the usage of online IM tactics in job-related contexts. Consequently, we believe that the results of this study will contribute to the academic and practitioner understanding of social media networks as a platform for optimizing IM tactics and affecting executive career success.

REFERENCES

Abbasi, A., and Chen, H. 2008. "CyberGate: A Design Framework and System for Text Analysis of Computer-Mediated Communication," *MIS Quarterly* (32:4), pp. 811-837.

Abbasi, A., Zhang, Z., Zimbra, D., Chen, H., and Nunamaker, J. F., Jr. 2010. "Detecting Fake Websites: The Contribution of Statistical Learning Theory," *MIS Quarterly* (34:3), pp. 435-461.

Abbasi, A., Albrecht, C., Vance, A., and Hansen, J. 2012. "Metafraud: A Meta-Learning Framework for Detecting Financial Fraud," *MIS Quarterly* (36:4), pp. 1293-1327.

Agarwal, A., Xie, B., Vovsha, I., Rambow, O. & Passonneau, R., 2011. "Sentiment Analysis of Twitter Data." *Proceedings of the Workshop on Languages in Social Media*, pp. 30-38.

Aggarwal, C. C., & Zhai, C. 2012. *Mining Text Data*. New York, NY: Springer.

Baker, G. P., and Hall, B. J. 2004. "CEO Incentives and Firm Size," *Journal of Labor Economics* (22:4), pp. 767–798.

Benthaus, 2014. "Making the Right Impression for Corporate Reputation: Analyzing Impression Management of Financial Institutions in Social Media," *Proceedings of European Conference on Information Systems (ECIS) 2014,* Tel Aviv, Israel, June 9-11, 2014.

Berman, S. L., Wicks, A. C., Kotha, S., and Jones, T. M. 1999. "Does Stakeholder Orientation Matter? The Relationship between Stakeholder Management Models and Firm Financial Performance," *Academy of Management Journal* (42:5), pp. 488-506.

Berry, M. W., Mohamed, A. H., & Yap, B. W. 2015. "Soft Computing in Data Science," *Proceedings of First International Conference, SCDS 2015* (545).

Bishop, C. M. 2006. *Pattern Recognition and Machine Learning*, New York, NY: Springer.

Bohnert, D., and Ross, W.H. 2010. "The Influence of Social Networking Websites on the Evaluation of Job Candidates," *Cyberpsychology, Behavior and Social Networking* (13:3), pp. 341-347.

Bolino, Mark C. 1999. "Citizenship and Impression Management: Good Soldiers or Good Actors?" *Academy of Management Review* (24:1), pp. 82-98.

Bolino, M., and Turnley, W. H. 1999. "Measuring Impression Management in Organizations: A Scale Development Based on the Jones and Pittman Taxonomy," *Organizational Research Methods* (2:2), pp. 187 -206.

Bolino, M. C. and Turnley, W. H. 2003. "More Than One Way to Make an Impression: Exploring Profiles of Impression Management," *Journal of Management* (29:2), pp. 141-160.

Bolino, M.C. Kacmar, K.M. Turnley, W.H., and Gilstrap, J.B. 2008. "A Multi-Level Review of Impression Management Motives and Behaviors," *Journal of Management* (34:6), pp. 1080–1109.

boyd, D. M. 2008. *Taken Out of Context: American Teen Sociality in Networked Publics*. Ph.D. Thesis. University of California Berkeley.

Boylu, F., Aytug, H., and Koehler, G. J. 2010. "Induction Over Strategic Agents," *Information Systems Research* (21:1), pp. 170-189.

Carter, M. and Grover, V. 2015. "Me, My Self, and I (T): Conceptualizing Information Technology Identity and its Implications," *MIS Quarterly* (39:4), pp. 931-958.

Carter, S. M. 2006. "The Interaction of Top Management Group, Stakeholder, and Situational Factors on Certain Corporate Reputation Management Activities," *Journal of Management Studies* (43:5), pp. 1145-1176.

42

Chun, R. 2005. "Corporate Reputation: Meaning and Measurement," *International Journal of Management Reviews* (7:2), pp. 91-109.

Conlon, D.E., and Murray, N.M. 1996. "Customer Perceptions of Corporate Responses to Product Complaints: The Role of Explanations," *The Academy of Management Journal* (39:4), pp.1040–1056.

Davies, G., and Miles, L. 1998. "Reputation Management: Theory versus Practice," *Corporate Reputation Review* (2:1), pp. 16-27.

Diga, M., and Kelleher, T. 2009. "Social Media Use, Perceptions of Decision-Making Power, and Public Relations Roles," *Public Relations Review* (35:4), pp. 440-442.

Duda Richard O. & Peter E. Hart. 1973. *Pattern Classification and Scene Analysis*, New York: Wiley.

Ellison, N. B., and boyd, d. 2013. *Sociality through Social Network Sites*. Oxford University Press, pp. 151-172.

Elsbach, K.D., Sutton, R.I., and Principe, K.E. 1998. "Averting Expected Challenges through Anticipatory Impression Management: A Study of Hospital Billing," *Organization Science* (9:1), pp. 68–86.

Fandt, P. M., and Ferris, G. R. 1990. "The Management of Information and Impressions: When Employees Behave Opportunistically," *Organizational Behavior and Human Decision Processes* (45:1), pp. 140-158.

Feldman, R. and Sanger, J., 2007. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press.

Frank, B.H. 2014. "How T-Mobile boss John Legere uses Twitter to Stay Ahead of Rivals", Retrieved Feb 29, 2016 from http://www.geekwire.com/2014/john-legere-says-twitter-account-key-t-mobiles-business/.

Gaines-Ross, L. 2000. "CEO Reputation: A Key Factor in Shareholder Value," *Corporate Reputation Review* (3:4), pp. 366-370.

Gardner, W. L., and Avolio, B. J. 1998. "The Charismatic Relationship: A Dramaturgical Perspective," *Academy of Management Review* (23:2), pp. 32-58.

Gattiker, U. E., and Larwood, L. 1988. "Predictors for Managers' Career Mobility, Success, and Satisfaction," *Human Relations* (41:8), pp. 569–591

Gilmore, D. C., and Ferris, G. R. 1989. "The Effects of Applicant Impression Management Tactics on Interviewer Judgments," *Journal of Management* (15:4), pp. 557-564.

Godfrey, J., Mather, P., and Ramsay, A. 2003. "Earnings and Impression Management in Financial Reports: The Case of CEO Changes," *Abacus* (39:1), pp. 95-123.

Goffman, E. 1959. *The Presentation of Self in Everyday Life*. Garden City, NY: Anchor Books.

Gordon, R. A. 1996. "Impact of Ingratiation on Judgments and Evaluations: A Meta-Analytic Investigation," *Journal of Personality and Social Psychology* (71:1), pp. 54.

Greene, W.H., 2003. *Econometric Analysis*, Upper Cliffs, NJ: Prentice Hall.

Gutteridge, T.G. 1973."Predicting Career Success of Graduate Business School Alumni," *Academy of Management Journal* (16:1), pp. 129–137.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, RL. 2006. *Multivariate Data Analysis*, Pearson Prentice Hall.

Han, J., Kamber, M., and Pei, J. 2011. *Data Mining: Concepts and Techniques: Concepts and Techniques*, Waltham, MA: Elsevier.

Harwood, T. G., and Garry, T. 2003. "An Overview of Content Analysis," *The Marketing Review* (3:4), pp. 479-498.

Heslin, P. A. 2005. "Conceptualizing and Evaluating Career Success," *Journal of Organizational Behavior* (26:2), pp. 113–136.

Hollander, S., Pronk, M., and Roelofsen, E. 2010. "Does Silence Speak? An Empirical Analysis of Disclosure Choices during Conference Calls," *Journal of Accounting Research* (48:3), pp. 531-563.

Hughes, E. C. 1937. "Institutional Office and the Person," *American Journal of Sociology* (43:3), pp. 404–413.

Jensen, M. C., and Murphy, K. J. 1990. "CEO Incentives—It's Not How Much You Pay, But How," *Harvard Business Review* (68:3), pp. 138-153.

Jones, E., and Pittman, T. 1982. *Toward a General Theory of Strategic Self-Presentation,* Hillsdale, NJ: Erlbaum.

IBM, 2012, "*Global CEO Study*". Retrieved from http://www-935.ibm.com/services/us/en/c-suite/ceostudy2012/infographic-01.html.

Kacmar, K. M., Delery, J. E., and Ferris, G. R. 1992. "Differential Effectiveness of Applicant Impression Management Tactics on Employment Interview Decisions," *Journal of Applied Social Psychology* (22:16), pp. 1250-1272.

Kane, G. C., Alavi, M., Labianca, G. J., and Borgatti, S. 2014. "What's Different about Social Media Networks? A Framework and Research Agenda," *MIS Quarterly* (38:1), pp. 274-304.

Kishore, R., Agrawal, M., and Rao, H. R. 2004. "Determinants of Sourcing during Technology Growth and Maturity: An Empirical Study of E-commerce Sourcing," *Journal of Management Information Systems* (21:3), pp. 47-82.

Krippendorff, K. 2012. *Content Analysis: An Introduction to its Methodology*, Thousand Oaks, CA: Sage.

Leary, M. R., and Kowalski, R. M. 1990. "Impression Management: A Literature Review and Two-Component Model," *Psychological Bulletin* (107:1), pp. 34-47.

Lee, P.M. and James, E.H., 2007. "She'-e-os: Gender Effects and Investor Reactions to the Announcements of Top Executive Appointments," *Strategic Management Journal* (28:3), pp. 227-241.

Lewis, D. D., & Catlett, J. 1994. "Heterogeneous Uncertainty Sampling for Supervised Learning," *Proceedings of ICML-94*, pp. 148–156.

Li, Y. M., and Shiu, Y. L. 2012. "A Diffusion Mechanism for Social Advertising over Microblogs," *Decision Support Systems* (54:1), pp. 9-22.

Liden, R. C., and Mitchell, T. R. 1988. "Ingratiatory Behavior in Organizational Settings," *Academy of Management Review* (13:4), pp. 572-587.

Longenecker, C. O., Sims, H. P., and Gioia, D. A. 1987. "Behind the Mask: The Politics of Employee Appraisal," *The Academy of Management Executive (*1:3), pp. 183-193.

Lucero, M., Kwang, A. T. T., and Pang, A. 2009. "Crisis Leadership: When Should the CEO Step Up?," *Corporate Communications: An International Journal* (14:3), pp. 234-248.

Luo, X., Zhang, J., and Duan, W. 2013. "Social Media and Firm Equity Value," *Information Systems Research* (24:1), pp. 146-163.

Mangold, W. G., and Faulds, D. J. 2009. "Social Media: The New Hybrid Element of the Promotion Mix," *Business Horizons* (52:4), pp. 357-365.

Marwick, A. E., and boyd, D. 2010. "I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience," *New Media and Society (*13:1), pp. 114-133.

McGuire, J.B., Sundgren, A. and Schneeweis, T., 1988. "Corporate Social Responsibility and Firm Financial Performance," *Academy of Management Journal* (31:4), pp. 854-872.

Mohamed, A. A., Gardner, W. L., and Paolillo, J. G. 1999. "A Taxonomy of Organizational Impression Management Tactics," *Advances in Competitiveness Research* (7:1), pp. 108-130.

Morris, R. 1994. "Computerized Content Analysis in Management Research: A Demonstration of Advantages and Limitations," *Journal of Management* (20:4), pp. 903-931.

Murphy, K.J. 1985. "Corporate Performance and Managerial Remuneration: An Empirical Analysis," *Journal of Accounting and Economics* (7:1), pp. 11-42.

Murphy, K. J. 1999. "Executive Compensation," *Handbook of Labor Economics* (Vol. 3B, pp. 2485-2563). Amsterdam, Netherlands: Elsevier Science.

Ng, H. T., Goh, W. B., and Low, K. L. 1997. "Feature Selection, Perceptron Learning, and a Usability Case Study for Text Categorization," *Proceedings of SIGIR-97*, 20th ACM International Conference on Research and Development in Information Retrieval, pp. 67–73.

Nicholson, N. 2000. "Motivation-Selection-Connection: An Evolutionary Model of Career Development," *Career Frontiers: New Concepts of Working Life*, pp. 54-75.

Oh, O., Agrawal, M., and Rao, H. R. 2013. "Community Intelligence and Social Media Services: A Rumor Theoretic Analysis of Tweets during Social Crises," *MIS Quarterly* (37:2), pp. 407-426.

O'Reilly, C.A., Doerr, B., Caldwell, D.F. and Chatman, J.A., 2014. "Narcissistic CEOs and Executive Compensation," *The Leadership Quarterly* (25:2), pp. 218-231.

Pant, G., and Srinivasan, P. 2013. "Status Locality on the Web: Implications for Building Focused Collections," *Information Systems Research* (24:3), pp. 802-821.

Parks, M. R. 2010. "Who Are Facebook Friends? Exploring the Composition of Facebook Friend Networks," *2010 Annual Conference of the International Communication Association, Singapore.*

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., and Duchesnay, É. 2011. "Scikit-learn: Machine Learning in Python," *The Journal of Machine Learning Research* (12), pp. 2825-2830.

Pollach, I., and Kerbler, E. 2011. "Appearing Competent A Study of Impression Management in US and European CEO Profiles," *Journal of Business Communication* (48:4), pp. 355-372.

Rettie, R. 2009. "Mobile Phone Communication: Extending Goffman to Mediated Interaction," *Sociology* (43:3), pp. 421-438.

Rogers, W., 1994. "Regression Standard Errors in Clustered Samples," *Stata Technical Bulletin* (*3*:13).

Rosenfeld, P. R., Giacalone, R. A., and Riordan, C. A. 1995. *Impression Management in Organizations: Theory, Measurement, and Practice*, New York: Routledge.

Schniederjans, D., Cao, E. S., and Schniederjans, M. 2013. "Enhancing Financial Performance with Social Media: An Impression Management Perspective," *Decision Support Systems* (55:4), pp. 911-918.

Schütze, H., Hull, D. A., and Pedersen, J. O. 1995. "A Comparison of Classifiers and Document Representations for the Routing Problem," *Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 229-237.

SEC. 2013. "*U.S. Securities and Exchange Commission Report on Investor Alert: Social Media and Investing -- Stock Rumors*". Retrieved from https://www.sec.gov/oiea/investor-alerts bulletins/ia_rumors.html.

SEC. 2014. "*U.S. Securities and Exchange Commission Report on Executive Compensation*". Retrieved from https://www.sec.gov/answers/execomp.htm.

Seibert S.E., Kraimer M.L., and Liden R.C. 2001. "A Social Capital Theory of Career Success," *Academy of Management Journal* (44:2), pp. 219–237.

Singh, V., and Vinnicombe, S. 2001. "Impression Management, Commitment and Gender: Managing Others' Good Opinions," *European Management Journal* (19:2), pp. 183-194.

Staw, B. M., and Epstein, L. D. 2000. "What Bandwagons Bring: Effects of Popular Management Techniques on Corporate Performance, Reputation, and CEO Pay," *Administrative Science Quarterly* (45:3), pp. 523-556.

Sutton, R. I., and Callahan, A. 1987. "The Stigma of Bankruptcy: Spoiled Organizational Image and Its Management," *Academy of Management Journal* (30), pp. 405– 436.

Tong, S. T., Van Der Heide, B., Langwell, L., and Walther, J. B. 2008. "Too Much of a Good Thing? The Relationship between Number of Friends and Interpersonal Impressions on Facebook," *Journal of Computer-Mediated Communication* (13:3), pp. 531–549.

Turnley, W. H., and Bolino, M. C. 2001. "Achieving Desired Images while Avoiding Undesired Images: Exploring The Role of Self-Monitoring in Impression Management," *Journal of Applied Psychology* (86:2), pp. 351-360.

Vapnik, V. N., & Chervonenkis, A. Y. 1971. "On the Uniform Convergence of Relative Frequencies of Events to their Probabilities," *Theory of Probability & Its Applications* (16:2), pp. 264-280.

Vonk, R. 1998. "The Slime Effect: Suspicion and Dislike of Likeable Behavior toward Superiors," *Journal of Personality and Social Psychology* (74:4), pp. 849-864.

Vonk, R. 2002. "Self-Serving Interpretations of Flattery: Why Ingratiation Works," *Journal of Personality and Social Psychology* (82:4), pp. 515-526.

Wade, J. B., Porac, J. F., Pollock, T. G., and Graffin, S. D. 2006. "The Burden of Celebrity: The Impact of CEO Certification Contests on CEO Pay and Performance," *Academy of Management Journal* (49:4), pp. 643-660.

Wayne, S. J., and Ferris, G. R. 1990. "Influence Tactics, Affect, and Exchange Quality in Supervisor-Subordinate Interactions: A Laboratory Experiment and Field Study," *Journal of Applied Psychology* (75:5), pp. 487.

Wayne, S. J., and Liden, R. C. 1995. "Effects of Impression Management on Performance Ratings: A Longitudinal Study," *Academy of Management Journal* (38:1), pp. 232-260.

Wayne, S. J., and Kacmar, K. M. 1991. "The Effects of Impression Management on the Performance Appraisal Process," *Organizational Behavior and Human Decision Processes* (48:1), pp. 70-88.

Weber Shandwick. 2014. "*Socializing Your CEO III: From Marginal to Mainstream*". Retrieved Feb 29, 2016 from http://www.webershandwick.com/uploads/news/files/socializing-your-ceo-iii-exec-summary.pdf.

Weisbach, Michael S. 1988. "Outside Directors and CEO Turnover," *Journal of Financial Economics* (20), pp. 431-460.

Welch, J., and Byrne, J. A. 2001. *Jack: Straight from the Gut,* New York, NY: Warner Books.

Westphal, J. D., and Stern, I. 2006. "The Other Pathway to the Boardroom: Interpersonal Influence Behavior as A Substitute for Elite Credentials and Majority Status in Obtaining Board Appointments*," Administrative Science Quarterly* (51:2), pp. 169-204.

Westphal, J. D., and Stern, I. 2007. "Flattery Will Get You Everywhere (Especially If You Are a Male Caucasian): How Ingratiation, Boardroom Behavior, and Demographic Minority Status Affect Additional Board Appointments at Us Companies," *Academy of Management Journal* (50:2), pp. 267-288.

Westphal, J. D., and Graebner, M. E. 2010. "A Matter of Appearances: How Corporate Leaders Manage the Impressions of Financial Analysts about the Conduct of Their Boards," *Academy of Management Journal* (53:1), pp. 15-44.

Westphal, J. D., and Deephouse. D. L. 2011. "Avoiding Bad Press: Interpersonal Influence in Relations between CEOS and Journalists and The Consequences for Press Reporting about Firms and Their Leadership," *Organization Science* (22:4), pp. 1061-1086.

White, H., 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," *Econometrica* (48:4) pp.817-838.

Yukl, G., and Falbe, C. M. 1990. "Influence Tactics and Objectives in Upward, Downward, and Lateral Influence Attempts," *Journal of Applied Psychology* (75:2), pp. 132–140.

Yukl, G., and Tracey, J. B. 1992. "Consequences of Influence Tactics Used with Subordinates, Peers, and the Boss," *Journal of Applied Psychology* (77:4), pp. 525–535.

Zhu, J.; Wang, H.; Yao, T.; and Tsou, B. K. 2008. "Active Learning with Sampling by Uncertainty and Density for Word Sense Disambiguation and Text Classification," *Proceedings of COLING'08,* pp. 1137–1144.

SUPPORTING INFORMATION

This section (i) describes the different machine learning-based classification procedures we employ to classify tweets from executives to generate our independent variables, namely, IM tactics, and (ii) provides comparative metrics that enable us to select the eventual text classification algorithm.

Algorithmic Classification Procedures

Algorithmic classification procedures can be broadly classified into two types — supervised and unsupervised learning. Both types of procedures can help define and explain phenomena that are captured by the dataset and can be used for predicting value of researcher-selected target attributes (in our case, target attributes are IM tactics), knowing the values of the relevant input attributes. In supervised learning, a given data set is typically partitioned into two: a training dataset with known category labels, and a testing data set. The training data set is provided as input to the algorithm, over several iterations. After each iteration, the generated categories, sometimes referred as 'target attributes', are updated using human coders. In contrast, unsupervised learning does not involve prior training, in predicting the target attributes of researcher interest. For instance, a typical unsupervised text mining algorithm can determine which terms or phrases in a given text dataset are related and can group them into clusters. Such a procedure can be helpful in discovering hidden topics embedded in complex data and provide an organized view of the data to facilitate decision making processes. Since supervised learning procedures involve training data and incorporate the knowledge of expert manual coders to ease the algorithmic component, the output of such procedures often tends to be more accurate than unsupervised models (Berry, Mohamed, & Yap, 2015).

We adopt supervised learning in this paper to identify IM tactics in Twitter feeds for two main reasons. First, we need the highest possible accuracy despite the large size of our data set. In a setting such as ours where we are dealing with large unstructured text data, the accuracy of unsupervised methods is likely to be poor, despite the ability of unsupervised methods to provide a simpler view for decision makers to understand. Second, the human coder effort for classifying a relatively small training dataset can be minimal, and yet enables us to accurately derive IM tactics. Hence, we apply a supervised learning procedure in our research. In the first stage, the training dataset for the learning algorithm is generated with manual inputs from human coders to accurately identify the IM strategies from a given subsample of tweets.

Having decided on the supervised learning approach, we next compare and contrast several alternative algorithmic procedures that are available at our disposal prior to selecting the best possible algorithm for our purpose. We investigate the accuracy of the following four procedures: 1-Frequency table based learning (e.g., Naïve Bayes), 2- Similarity distance based learning (e.g., K-Nearest Neighbors), 3- Machine learning (e.g., Support Vector Machines), and 4- Neural network (e.g., Perceptron). A comparison chart of alternative approaches and their accuracy rates on our trained data set is shown in Table 1.A-1.

As we observe from Table 1.10, recently-proposed learning methods, such as SVM and Perceptron, have outperformed other statistical methods in similar classification tasks (Bishop, 2006). Theoretical foundations of support vector machine (SVM) algorithms trace back to Vapnik and Chervonenkis (1971). The intuition is as follows: SVM starts with the specification of known outcomes (e.g. IM tactics) and input attributes. Either these known outcomes may be binary (e.g. positive/negative) or multi-class attributes (e.g. customer segments).

Table 1.10: Comparison of Text Classification Algorithms			
Classification **Method**	Accuracy	Objective	Strength/Weakness
Perceptron (number of iterations=50)	**0.709**	Finds a hyperplane that separates the classes by adjusting weights and bias values of inputs	+ Efficient at learning the most important features from various data structures. - Due to high number of parameters, it may fail to find global optimum if training set is small.
Linear-SVM	0.664	Constructs a hyperplane that maximizes the margin between classes by using distance vectors	+ SVMs require less time to find global optimums and need less memory to store the predictive model. - Since the algorithm focuses on best isolation of classes, inter-related inputs may decrease the classification performance.
K-Nearest Neighbors Classifier (number of neighbors=15)	0.636	Classifies an attribute by using majority vote of its nearest neighbors based on Euclidian distance measure	+ Provides efficient solutions for low dimensional data types by using simple similarity distance functions. - Instead of learning from training set, this algorithm just uses the training set, thus less generalizable for independent data sets.
Naïve Bayes	0.611	Finds the likelihood of an attribute's class by using prior and posterior probabilities of given class	+ This frequency table based classifier is less parametric and provides efficient solutions for balanced data sets. - When there is interaction among inputs, algorithm fails to provide accurate results.

SVM then projects the input data into a multi-dimensional space, and then constructs a decision surface to maximize the margin (separation) between distinct classes. When we provide a new

tweet as an input to the procedure, the algorithm decides the class to which the tweet belongs based on distance between classes.

In contrast, Perceptron is a relatively more recent classification algorithm which falls under the broad category of neural network text classifiers. Perceptron classifier has successfully been applied to text classification problems (Schutze et al. 1995; Ng et al. 1997). The main intuition is as follows: we parse each tweet into a set of tokens, phrases or collection of words, with random weights assigned to the tokens. This set of tokens and weights comprises the perceptron layer, which the algorithm maintains and updates. Alternatively stated, an input set is denoted by a vector which is drawn from a lexicon of words in our text data, and then the set of weights is used to compute a function of inputs to arrive at a classification label for each observation (Aggarwal & Zhai, 2012). The Perceptron classifier checks the training dataset one observation at a time to predict their label (IM tactic) based on the observed inputs and weights. If the prediction is correct, iterations are continued. Otherwise, the observation is used to correct the set of weights. We employ a variation of this procedure to accommodate our multiclass classification problem.

Impression Management Tactics Training Document

Dimensions with Examples	Impression Management Tactics
Self-Promotion *If you need to get better organized and build systems, do what I do*	- Make people aware of your accomplishments. - Try to make a positive event that you are responsible for appear better than it actually is. - Try to take responsibility for positive events, even when you are not solely responsible. - Try to make a negative event that you are responsible for appear less severe than it actually is - Display your diplomas and/or awards that you have received. - Let others know that you have a reputation for being competent in a particular area. - Make public your talents or qualifications. - Declare that you have other opportunities outside your current job. - Talk about important people that you know. - Try to distance yourself from negative events that you were a part of. - Talk proudly about your experience or education. - Make people aware of your talents or qualifications. - Let others know that you are valuable to the organization. - Make people aware of your accomplishments.

Dimensions with Examples	Impression Management Tactics
Ingratiation *@xxx: I'm with Arie..excited to hang out with you tomorrow:)*	• Praise people for their accomplishments. • Do personal favors for people. • Offer to do something for someone that you are not required to do. • Compliment people on their dress or appearance. • Agree with a person's major ideas or beliefs. • Take an interest in a coworker's, board members or other stakeholders' personal life. • Imitate others' behavior or manner. • Spend time listening to people's personal problems even if you have little interest in them. • Compliment stakeholders so they will see you as likeable. • Take an interest in stakeholders' personal lives to show them that you are friendly. • Praise your colleagues for their accomplishments so they will consider you a nice person. • Use flattery and favors to make your colleagues, board members, peer directors, and even journalists like you more. • Do personal favors for stakeholders to show them that you are friendly.
Exemplification *the best thing you can learn is to forget yourself and serve the community*	• Try to highlight that you are a socially and environmentally responsible person. • Engage in social responsibility activities • Try to appear like a hard-working and dedicated manager, • Volunteer to help whenever there is the opportunity. • Make sure you are never seen wasting time.
Intimidation *Being rich should not allow you to treat people like sh&*!! @xxx* *You service is absolutely terrible - get on the ball! @xxx*	• Try to appear unapproachable or distant. • Make people aware that you can control things that matter to them. • Look intimidating to stakeholders when it will help you get your problem solved. • Let others know that you can make things difficult for them if they push you too far. • Use intimidation to get stakeholders to behave appropriately. • Deal strongly or aggressively with third-parties who interfere in your business. • Show stakeholders that you are powerful and competent enough to punish people.

Dimensions with Examples	Impression Management Tactics
Supplication *Ive tried for 13 yrs to fix the officiating in this league and I have failed miserably. Any Suggestions? I need help*	• Try to gain assistance or sympathy from people by appearing needy in some area. • Act like you know less than you do so people will help you out. • Advertise your incompetence in a particular area or about a particular issue. • Pretend to not understand something that you do understand. • Ask for help or assistance that you really do not need. • Try to appear helpless or needy. • Downplay your accomplishments. • Let others win arguments. • Try to agree with people even when you might disagree.

To arrive at an unbiased training dataset, we need to obtain a balanced number of tweets from each IM category. The research team will manually code the tweets from the training dataset. The coded training dataset are input used in the supervised learning algorithms. In order to ensure that the training set is unbiased, we employ a sampling-by-clustering approach to achieve this objective.

Sampling by Clustering

A typical machine learning problem is to maintain accurate classifiers while minimizing the number of observations to be coded. Traditional approaches to select observations included in training data sets include random sampling, stratified sampling, uncertainty sampling (Lewis and Catlett, 1994). Random and stratified samplings are traditional methods to generate an initial training set from the whole unlabeled corpus or body of text. However, these sampling techniques may not generate a representative training data set because the size of initial draw is generally

small and might include a large amount of unrelated text. In our setting, we need to ensure that our initial training set covers sufficient amount of tweets from each of the five IM strategies. Uncertainty sampling is another technique based on selecting obvious instances for labeling for which the learner is most certain (Lewis and Catlett, 1994). However, one concern with choosing this method is that selecting obvious instances for labeling is prone to selecting outliers as representative of the IM tactic.

We follow the sampling-by-clustering algorithm proposed by Zhu et al. (2008) to form initial data set. The sampling-by-clustering algorithm overcomes the problem of selecting representative samples. In summary, the entire unlabeled corpus of tweets is partitioned into a predefined number of clusters. Determining the number of clusters in this method primarily depends on the number of dimensions of interest and the inter-dependency among clusters. In our analysis, we eventually arrive at 20 clusters after examining cluster centroid distances when employing 5, 10, 15, and 20 clusters. We stop at 20 clusters since the relative change of the total distortion among clusters becomes small after this threshold. The sampling-by-clustering algorithm uses cosine-based distance measure and K-means clustering to estimate similarity among tweets and assign the subsequent tweet closest to the centroid of each cluster (See Zhu et al, 2008; Duda & Hart, 1973 for the details of clustering algorithm). Upon clustering the whole corpus, we next randomly select 200 tweets from each cluster totaling 4000 messages for manual coding.

Applying Customized Feature Extraction and Machine Learning to Predict IM Tactics

Our purpose in applying machine learning techniques suitable for unstructured textual data is to transform text-based tweets into a numerical feature matrix. Alternatively stated, we form a two-dimensional matrix where numerical columns represent features (of the tweet) and rows

represents observations (tweets). To begin our effort in predicting IM tactics, we attempt to efficiently assign a value to a feature for corresponding observation. Once we achieve a rectangular matrix represented with numerical values, we can next apply an appropriate machine *learning* technique to our data set. A widely accepted way to transform unstructured text into a rectangular matrix is to count the frequency of the words in the text using the "bag of words" approach. In our approach, a tweet is represented as a bag of words, a collection of words with no regard to the order in which each word occurs, and it is used to generate a vector of frequency counts of words for computational purposes (e.g. Pedregosa et al, 2011). Each unique word in a tweet represents a feature including stop words such as 'the', 'an', and 'in'. In the following example, a raw tweet (A) is transformed into (B):

(A) #We are live at 7 pm! Join @Gabby

(B) ["#", "We", "are", "live", "at", "7", "pm", "!", "Join", "@", "Gabby"]

However, when one considers the number of unique words and terms in English, our above approach to transformation will result in an extremely large feature matrix for a given text document. Fortunately, the scikit-learn module of python programming language enables us to address this problem by helping convert each tweet into a list of refined words or *n-grams*. For instance, the following example illustrates how different terms can be amalgamated into one dimension (Agarwal et al, 2011). Here, instead of including each term as one dimension, terms can be customized by binning them into similar meta strings and then amalgamated under one dimension (e.g., positive emoticons).

positive_emoticons <= [':)', ':-)', ' :)', ':D' ,'=)', ' : D ', '(:' , '(=']

For impression management (IM) tactics prediction, we adopt a similar approach and bin words into similar meta-strings so as to identify the tactic employed by the IM actor. For example, ingratiation is a tactic which focuses on showing flattery or opinion conformity directed at a specific audience in order to gain favor. A classical machine learning approach which employs a non-customizable feature extraction model will fail to differentiate the intention in the following example.

[1] A big thanks to @DavidKirkpatric for bringing together great minds to consider the interplay of technology and mankind! Very good stuff!"

[2] "I am late, thanks to Manhattan traffic. pic.twitter.com/pLkPLsE"

In the first tweet, a classical approach will be able to identify an ingratiation tactic towards @DavidKirkpatrick. However, in the second tweet, there is no ingratiation tactic being employed although actor uses the same phrase. Hence, we need to customize the tokenization process (of breaking down the tweets into weighted features) to form a reasonable-sized (in a computational sense) and reliable feature matrix. We attempt to separate audience-directed phrases from similar word combinations to remove such ambiguity. In this case, we specifically focused on commonly-used ingratiatory phrases and combined them with directed phrases such as "@XXX," implying that the message *mentions* a twitter account owner. We next amalgamated all the directed accounts under "@mention" tag and binned that tag with the common ingratiatory phrases as in the following example.

Ingratiation <= ['proud of @mention', 'love @mention', 'like @mention', 'excited for @mention', 'congratulations to @mention', 'thanks to @mention', 'well done @mention', 'well deserved @mention', 'good job @mention', 'happy for @mention'…]

Such syntactic ambiguity will cause an incorrect interpretation by linear classification algorithms, in addition to the problems that stem from the semantic ambiguity at the level of individual words. For instance, supplication can be represented by individual words such as "help", "need", and "please". Once manual coders label the tweets for this tactic, a classical algorithm will automatically focus on those unique words in the prediction stage. However, usage of these words in tweets may not always indicate supplication. For example, while "I need help" may signal a supplication intention, "you need help?" would actually imply the opposite. Thus, we apply bi-grams, tri-grams, and unique word combinations into the feature extraction process to remove such ambiguity. As a result of these steps, linear classifiers, especially perceptron, would be able to perform efficiently and increase the overall prediction accuracy.

The code for this study will be provided upon request.

AN ANALYTIC VIEW OF THE EXECUTIVE REPUTATION IN SOCIAL MEDIA NETWORKS AS A DECISION MECHANISM

ABSTRACT

This paper investigates the external constituents of executive reputation and identifies their consequences from a signaling theoretical lens. We examine how externally established executive reputation, in the form of word-of-mouth, may affect managerial survivability. The recent advances in information technology require corporate executives to manage and monitor their personal reputation in the eye of internal and external stakeholders and shareholders on various communication platforms. We aim to develop theoretical and empirical support for the concept that external cues of executive reputation lead to consequences for a top executive's career path. In this paper, we first analyze the credibility of online word-of-mouth and fundamental differences of social media networks from traditional news media as an information source, and then provide detailed dimensions of the reputation concept formed through separate information sources. Second, we illustrate the outcomes of executive reputation formed in social media networks while focusing on the consequences of such reputation in the corporate world. Finally, we apply multiple data mining techniques to quantify the effects between executive reputation and managerial survivability.

Keywords: word-of-mouth, executive reputation, survivability, data mining.

INTRODUCTION

"Regard your good name as the richest jewel you can possibly be possessed of - for credit is like fire; when once you have kindled it you may easily preserve it, but if you once extinguish it, you will find it an arduous task to rekindle it again. The way to gain a good reputation is to endeavor to be what you desire to appear."

Socrates *Greek philosopher in Athens (469 BC - 399 BC)*

This paper explores the impact of an important intangible asset, executive reputation, on firm and executive level outcomes. Reputation is defined as favorable/unfavorable public opinions about an individual or an organization and is seen as a strategic component of the firm (Penrose 1959; Wernerfelt, 1984) and one of the most valuable intangible assets (Hall, 1992). Prior findings confirm its value by demonstrating that both organizations and top management team engage in strategies to avoid negative reputation in media by influencing journalists and other press (e.g. Westphal & Deephouse, 2011). In addition, researchers have examined the consequences of reputation, such as financial performance and executive compensation, using proxies including volume and valence of coverage in news media to represent the reputation construct (e.g. Deephouse, 2000). Aside from previously analyzed news media resources, the growing plethora of social media networks (SMN) and their impacts on executive reputation pose an important question to be examined. In this study, we investigate the executive reputation construct especially in current popular communication and information platforms and its consequences from a signaling theory perspective.

SMN such as Facebook and Twitter are not only channels for disseminating business news for corporations but also an arena for participation in which top executives interact with the public. SMN are bi-directional communication channels where stakeholders and other interested parties can interact with each other. These interactive platforms contain news, word-of-mouth, and

external assessments, which serve as key constituents that invoke reputation that is an intangible asset in the corporate world (Van Hoye & Lievens, 2009). In other words, both internal and external stakeholders can actively participate in the ongoing process of influencing assessments of corporations and their executives by generating word-of-mouth based on perceptions and information gathered from various resources.

Stakeholders express their opinions as they search for information, gain knowledge, and make interpretations based on news and actions about an organization and its executive team. Once they have formed an opinion, they share it with others and their personal perceptions become public (Mills, 1959). Propagation of perceptions is indeed important here because it eventually shapes the overall external assessment about a focal figure (Heath, 1996), the top executive. In contrast to traditional press media, SMN aggregate various assessments about a focal firm and its executive team from different stakeholder groups including customers, analysts, communities, prospective employees, agencies, and investors.

Assessments provided by stakeholders about the focal firm and its executive team prioritize the relevance of certain attributes of the executive reputation more noticeable than others. For instance, while current and potential customers may be more concerned about an executive's capability for producing high quality products or services, current and prospective employees might care about trustworthiness and share their opinions regarding trust. Likewise, while analysts and investors may be more interested in the role of an executive on financial performance of the focal firm, competitors within the same industry may post about deficiencies pertaining to existing capabilities of the executives and focal firm (Mishina, Block, & Mannor, 2012). Thus, SMN coverage includes a wide spectrum of information aggregated from various stakeholders and plays

an extended social arbiter role, which may have implications on reputation (Deephouse, 2000; Pollock, & Rindova, 2003).

Top executives are increasingly recognized as high profile figureheads for their firms and firm reputation is highly influenced from the public's views of a company's top leadership (Arvidsson, 2006). Public opinions and assessments about top executives represent how their actions and behaviors are perceived by the world and eventually form a reputation of them. Ideally, a favorable stakeholder perception about an executive can create value for their companies and their professional benefits (Rein et al., 2006). Thus, the influence of executive reputation is positive, as in the case of company shareholder perceptions, where the media can play an important role in reflecting such reputation (Arvidsson, 2006).

SMN adds another layer on top of reputation strategies, as SMN offer stakeholders an option to have a literal voice that speaks directly to firm and other stakeholders in everyday conversations. Interaction among stakeholders through social media is persistent and visible to all. Social media serves as a mirror to reflect "public displays of connections" (Boyd & Ellison, 2010). In addition, SMN offers exponential spread of content, along with its unprecedented levels of accessibility. These features of SMN make it the fastest-growing reputation management channel in the world (Evans, 2012).

In a nutshell, SMN have the effect of mirroring a collective opinion apart from conventional media. Users search for information, gain knowledge, and make interpretations based on communication about an organization and its leaders. Once they have built an image, they share it with others and the personal subjective opinion turns into a collective opinion about what an organization's management team is and what it should be. Therefore, SMN are now considered a

62

valuable and credible source for distinct stakeholders whose perceptions have implications on firm decisions (Aggarwal & Singh, 2013).

In this study, we aim to contribute to the information systems literature by investigating executive reputation in SMN and comparing our findings with prior research by integrating the value of public opinions in our model. We argue that, although previously omitted, a manager's personal reputation in SMN can affect organizational perception and executive success. In addition, we tease out the effects of distinct groups -stakeholders and shareholders- on managerial level outcomes. Specifically, we seek to answer two questions. First, does a general public reputation in SMN for an executive have an effect on managerial survivability in current position? Second, which specific dimensions of such reputation influence maintaining and extending top executive positions within the firm? We draw on signaling theory to explain how this reputation can be utilized as a valuable asset to extend managerial survivability.

BACKGROUND

This section is designed as follows. First, we review the literature to explain executive reputation across two different media platforms. Second, we narrow down the research question to SMN and emphasize organizational and managerial perspectives.

Conventional Media vs. Social Media Networks

Previous research explored executive reputation mainly in the conventional news media setting in the forms of broadcast news and print media. Conventional media generally serve as an information provider and aim to reduce information asymmetry between firms and stakeholders.

Some stakeholders lack direct experience with a firm. Instead they depend on information intermediaries, such as the government, rating agencies, and news media, who, "screen, spin, and broker information for us; they help us make sense of companies' complex activities – and so affect company reputation" (Fombrun, 1996). Previous studies provide both theoretical and empirical evidence that media shapes the way stakeholders assess and interpret information about firms by framing explanations in positive or negative phrases (McCombs, Llamas, Lopez-Escobar, & Rey, 1997). However, today SMN are attracting new audiences and making both internal and external stakeholders aware of company or managerial events including prospective employees and potential customers. SMN allow a user to generate content in real time and make that content available to a wide audience immediately. The stakeholders' perceptions are formed in a short period due to the easy and efficient information spread about firms and executives. More importantly, SMN provide platforms for a vast range of stakeholders to share and exchange their judgments, even about the news circulating in conventional media, whereas conventional media involves limitations in terms of interactivity by primarily reaching target audience via mass broadcast communication.

Managing one's reputation in conventional news media requires special effort. Recent research has focused on efforts to avoid bad press and manage impressions of stakeholders (Hayward, Rindova, & Pollock, 2004; Westphal & Deephouse, 2011; Westphal et al, 2012). In addition, in search of reliable information, stakeholders give credit to theoretically weak executive ratings and certification contest rankings in order to evaluate top management performance (e.g. Wade, Porac, Pollock, & Graffin, 2006). However, with the rapid growth of online networking outlets, SMN are leveraged as credible information sources for issues ranging from social

movements (Oh, Agrawal, & Rao, 2013), and political decisions (Kushin & Yamamoto, 2010) to financial markets (Bollen et al., 2011).

Users can export traditional news feeds by sharing them through SMN, while traditional media can also integrate social media channels into its news practices, which forms a dyadic relation between traditional media and SMN (Tufekci & Wilson, 2012). Thus, news media and SMN should not be regarded as mutually exclusive. As positive reputation brings more negotiation power, prestige, celebrity status, and higher compensation packages to top executives (Milbourne, 2003; Hayward, Rindova, & Pollock, 2004; Wade et. al, 2006), a positive reputation in SMN - containing not only news but also recipients' interpretation can confer a valuable strategic benefit for top managers.

Reputation in Social Media Networks (Corporate vs. Executive Lens)

From a corporate view, reputation is a valuable intangible asset for a firm (Hall, 1992). Although SMN had been first perceived as a formidable tool for reputation management, today it has become an inevitable platform, an IT artifact, for improving public reputation (Berger, Klier, and Probst, 2014). Word-of-mouth (Van Hoye & Lievens, 2009) and media coverage (Fombrun & Shanley, 1990) are considered external assessments that induce the executive reputation. Instead of using solely conventional press coverage as an external cue for reputation, we expand the information base and utilize stakeholders' news sharing efforts and their interpretations(2) in the form of electronic word-of-mouth as comprehensive indicators to analyze reputation construct.

---

2 Shareholders (with a financial stake in the firm) and stakeholders (interested in firm performance other than just financial performance) often overlap in strategy and finance literature with stakeholders being a more inclusive term. In this chapter, we consider the distinct positioning of stakeholders as

From an executive point of view, previous research posits that an executive's favorable media reputation has a positive effect both on his/her compensation (Milbourn, 2003) and on corporate activities (Francis et al. 2008). Reputation in media not only influences stakeholder perceptions of firms but also the reputation of corporate leaders (Wade et al. 2006). Positive media reputation can strengthen managerial authority and enhance executive discretion over corporate policy while also increasing their career prospects (Hayward, Rindova, & Pollock, 2004). In contrast, negative media reputation can reduce managerial power and harm a top executive's image, thus diminishing manager discretion over corporate policy and damaging career prospects (Wiesenfeld, Wurthmann, & Hambrick, 2008). Since past studies report findings regarding executives' reputation mainly in traditional news media settings (which mainly lack additional commentary of a wide variety of stakeholders), we ask if the results can be generalized once we incorporate such public opinions, judgements, and assessments into the analysis. We specifically analyze the relationships between the executives' online reputation and managerial level outcomes. These relationships are depicted below in Figure 2.1.

---

entities that behave in a socially conscious manner. This differentiation is important as we attempt to understand the semantic content, context, and the intended audience of social media messages.
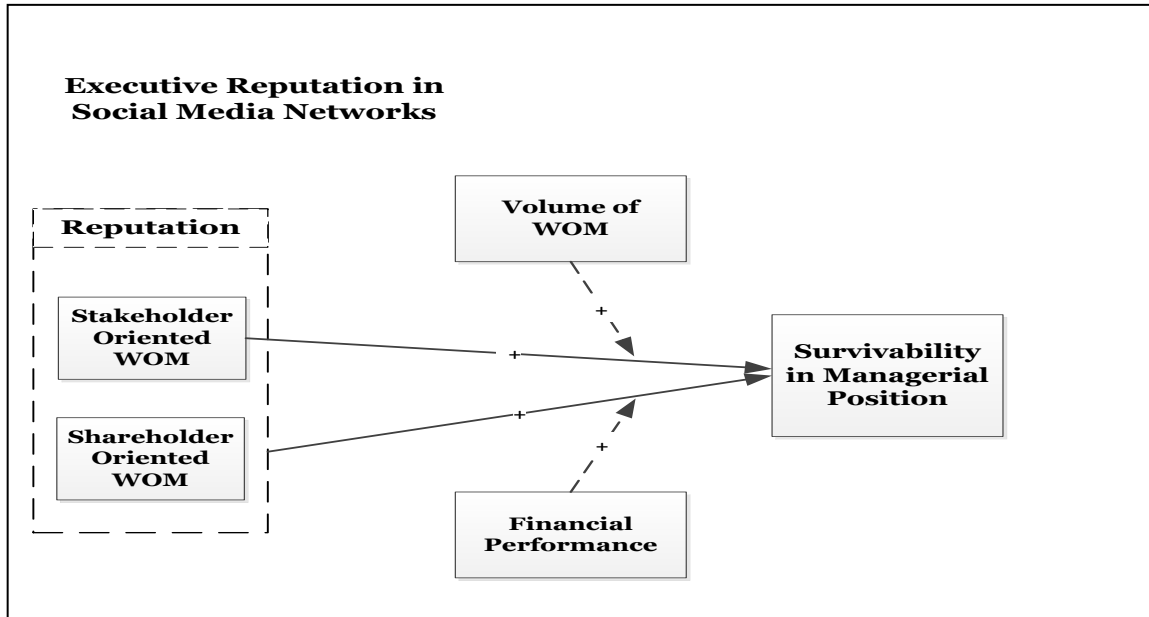
Figure 2.1. Effects of executive reputation in social media networks

THEORETICAL MODEL

Executive Reputation in Social Media Networks

The executive reputation construct serves as the focal point of our model and is defined as "favorable/unfavorable opinions about an individual in a social network" by Lin (1999). Word-of-mouth is considered as a key external assessment that induces executive reputation construct that eventually affects managerial level outcomes (Van Hoye & Lievens, 2009). In this context, the type of social network is a substantial determinant of word-of-mouth which subsequently determines the spread of reputation in the focal network. The recipient-source framework (Gilly et al., 1998) has mainly guided research on other determinants of word-of-mouth and previous studies have largely focused on the effect of two dimensions—valence and volume —of word-of-mouth in various research contexts (e.g. Aggarwal et al., 2012). Thus, in our study, we will first explain the valence and volume dimensions of word-of-mouth and then illustrate key facts when

67

word-of-mouth occurs in a SMN type network structure, namely electronic word-of-mouth. Finally, we will utilize signaling theory to examine the construct relationships in our model.

Valence of Word-of-mouth in Social Media Networks

Valence is defined as the degree that content of overall coverage is positive or negative. Previous evidence shows that the valence of coverage in SMN shapes the stakeholders' perception (e.g. Aggarwal et al., 2012). Distinct audience groups can generate word-of-mouth by framing the content of information through their interpretations and presentations (Gitlin, 1980). Stakeholder's perception of an executive's action depends on the positive or negative framing of the content of information (Westphal & Deephouse, 2011). This perception has implications on firm practices such as executive power and reputation influence the executives' tenure in their positions (Fredrickson et al, 1988). Therefore, disseminated information about an executive may affect the processes of decision making through the valence of the coverage.

The degree of public support for an executive, either favorable or unfavorable, may have implications on executive reputation. Previous studies provide both theoretical and empirical evidence that coverage in several media platforms shapes the way stakeholders assess and interpret information about firms by framing explanations in positive or negative phrases (McCombs et al, 1997). Such coverage not only influences stakeholder perceptions of firms but also affects the reputation of corporate leaders (Wade et al. 2006). However, to the best of our knowledge, the effects of executive reputation in the form of word-of-mouth have not been examined in the electronic realm of SMN such as popular sites Twitter and Facebook.

Electronic word-of-mouth in social media is specified as any statements users share via online interactions (Kietzmann et al. 2012). For example, users share their opinions as short messages on the social media site, Twitter or Facebook. As credible information sources, Twitter and Facebook posts are also heavily examined as forms of word-of-mouth in a variety of contexts such as emergency management (Oh, Agrawal, & Rao, 2013), political decisions (Kushin & Yamamoto, 2010), financial markets (Bollen et al, 2011), and branding strategies (e.g. Jansen et al, 2009). Although such studies provide evidence about the influence of word-of-mouth on various organizational practices, we are interested in teasing out the content about executives' reputations by analyzing the valence of word-of-mouth and its effect on managerial survivability.

Effects on Managerial Survivability

SMN allow users to generate content in real time. The stakeholders' perceptions are formed in a short period due to the efficient information spread about executives in SMN. Organizations and individuals are using SMN as a way to enhance, sustain, and defend their reputation (Mangold & Faulds, 2009; Li & Shiu, 2012). Thereby, the content of information in SMN is a credible information source for stakeholders (Aggarwal et al., 2012) including decision makers who influence executive careers.

We draw on signaling theory (Spence, 1973) to explain the effect of word-of-mouth on managerial career decisions. Signals are observable attributes of a firm and top management team that can change the perception of stakeholders (Sanders & Boivie, 2004). Signals need not be broadcasted only by firms' internal channels; rather, any external monitor such as media and social media can emanate signals for executive actions (Fombrun & Shanley 1990). Spence (1973) stated that signals should be cost associated and observable. Since writing a post on SMN

about an executive and sharing a personal assessment cost time and effort, it satisfies the criterion of cost for an efficacious signal (Aggarwal et al., 2012). Since a high volume of comments and word-of-mouth about an executive is likely to be monitored more readily by an observer (e.g. shareholders and board of directors), it also satisfies the criterion of observability for an efficacious signal. Therefore, valence of word-of-mouth in SMN should act as a signal to key stakeholders and shareholders about the public reputation of an executive.

Signals generated and spread through word-of-mouth that become contagious in SMN will provide ease of access to information about executives for stakeholders and shareholders. Such viral events may even cause an alert for oblivious parties through SMN where connectivity and interaction play key roles (Sedereviciute & Valentini, 2011). On one hand, information transferred through SMN will allow for rapid transmission of innovations, promotions, expertise, and best practices between firm and stakeholders, which are strategically valued by the firms in subsequent phases (Geletkanycz, Boyd, & Finkelstein, 2001). On the other hand, negative opinions based on misconduct or an unfortunate statement may seriously harm the public reputation of an executive and lead negative outcomes for executives' careers. A social media post that became contagious in 2013 about Abercrombie Fitch CEO, Mike Jeffries, offended overweight customers.

> @SofiaJasmine 23 Aug 2013
> #misslawrence on hateful #abercrombie CEO #MikeJeffries: "He looks like a big piece of provolone cheese" #QuoteOfTheDay

A collective negative word-of-mouth about Mike Jeffries was fostered by the vast network structure of SMN and resulted in the public disgrace of the Abercrombie Fitch CEO.

Social network tie structures in SMN can be an important means to disseminate reputation cues to others such that these ties are like prisms and can deliver information to others about the focal individual (Podolny, 2001). Marwick (2013) emphasized how the social network structure of SMN and electronic word-of-mouth transform the reputation management in the social media age. In the executive context, Mehra et al. (2006) proposed that the executives' reputation is associated with the external social influence. Although sometimes a collective favorable reputation in the public domain may generate a positive perception and help executives open doors to the public sector, push sales and attract new contracts (Agrawal and Knoeber, 2001; Kirchmaier and Stathopoulos, 2008), a collective negative word-of-mouth in dynamic social platforms may also influence decision making regarding the careers of top executives. This can also exert additional pressure on firms to make changes in corporate governance that increase the risk of executive dismissal (Core et al. 2008). This leads us to our first hypothesis:

H1. Executive reputation established via social media networks is associated with managerial survivability such that overall favorable reputation will increase the survivability in managerial positions and ceteris paribus.

Sub-dimensions of Executive Reputation

Stakeholder-Oriented Reputation

Since SMN have become a core part of social interactions (Kane et al. 2014) where users interact with others and create social networks with no cost, companies have been extensively leveraging this fact to foster their reputation. Given that millions of people use online social media platforms, participants in these vast networks have the potential to influence large audiences at once (Schniederjans, Cao, & Schniederjans, 2013). In addition to providing considerable networking potential, these platforms markedly enhance communication among

both internal and external stakeholders.

In a stakeholder-oriented business culture, a broad spectrum of stakeholders are perceived by society as possessing a legitimate interest in corporate activities (Dhaliwal et al. 2012). Accordingly, we define stakeholder-oriented reputation as the valence of word-of-mouth among distinct audiences who may possess a legitimate interest in corporate activities such as community, customers, employees, environmentalists, and government. While SMN will boost the spread of word-of-mouth among a wide range of stakeholder groups, the valence of such reputation of an executive within stakeholder-oriented discussions will simultaneously affect the perception of the focal company.

Theoretical studies on the stakeholder literature suggest that the executives who are exclusively monitored with a stakeholder orientation would enhance the positive perception of the company (Donaldson & Preston, 1995). Specifically, consistent stakeholder oriented actions espoused by executives will trigger positive perceptions among key stakeholders who value such orientation (e.g. customers, employees and even prospective employees). For example, word-of-mouth about an executive's social and environmental responsibility action may be perceived as favorable by the community (e.g. environmental activists) and may strengthen the stakeholders' emotional attachment to the focal firm. Since chief executives and company reputation are found to be inextricably linked (Gaines-Ross, 2000), we propose that the valence of stakeholder oriented word-of-mouth about executives in social media networks will influence the corporate governance decisions. Thus;

H2a. Executive reputation established via social media networks is associated with managerial survivability such that specific stakeholder-oriented reputation will increase the survivability in managerial positions.

Although we have argued that the valence of stakeholder-oriented reputation is likely to send a signal to board members for a decision about an executive, the strength of this signal may depend on how big this signal is. Studies in socio-cognitive fields suggest that the amount of available information also influences people's perceptions (Fiske & Taylor, 1991). Prior studies show that media platforms legitimize entities by guiding public attention to those selected for coverage, therefore increasing the public's exposure to them (Kosicki, 1993). The degree of audience exposure through media is a function of the volume of the coverage an entity receives. Such exposure influences socio-cognitive processes germane to comprehension and liking (Pollock & Rindova, 2003). A vital mechanism through which high positive exposure in media results in favorable perceptions in the eye of audiences is that while audiences are unaware of their familiarity with a stimulus, they nevertheless demonstrate preference towards a stimulus to which they have been exposed more frequently (Pollock & Rindova, 2003). Briefly, social cognition studies suggest that repeated positive exposure to an object increases familiarity and leads to preference for the object (Zajonc, 1968).

In line with the aforementioned theoretical arguments, previous research in related contexts posit that the volume of word-of-mouth in SMN will increase the visibility of entities and most likely attract more observers to monitor their actions (Aggarwal and Singh, 2013; Aggarwal et al, 2012). Social media platforms such as Facebook and Twitter have no costs for user accounts and do not restrict information sharing. Thus, an increase in the volume of word-of-mouth will generate the following processes: 1- Volume of coverage will strengthen the position of the executives in their own social networks (Kietzmann et al, 2011; Burt, 1992). 2- Volume of coverage will increase the visibility and exposure of executives and the brands they are representing to audiences (Pollock & Rindova, 2003). 3- Increased awareness and familiarity

with an executive will facilitate favorable perception formation among stakeholders (Pollock & Rindova, 2003). Accordingly, we propose that there will be supplemental interaction such that the volume of word-of-mouth will moderate the relationship between stakeholder-oriented reputation of executives in social media networks and decision about their careers.

H2b. The effect of stakeholder-oriented executive reputation in SMN on managerial survivability will be stronger when the volume of word-of-mouth is high.

Shareholder-Oriented Reputation

In a shareholder-oriented business culture, companies give priority to shareholder value, while providing less legitimacy in affecting corporate activities and performance to other stakeholder groups (Dhaliwal et al. 2012). Accordingly, we define shareholder-oriented reputation as the valence of word-of-mouth among the audience who are interested in and concerned with the company's financial performance such as analysts, shareholders, and potential investors. When an audience (e.g. analysts) shares information about an executive's action, SMN boosts the spread of word-of-mouth among financial interest groups. The valence of executive reputation within shareholder-oriented discussions will simultaneously affect the perception of company's performance.

As a specific case, top executives' themselves can serve as a trigger for shareholder-oriented discussions in SMN. In general, individuals use a mental calculation to estimate the risk-benefit ratio of making private information public. In the case of top executives, a variety of risks such as inadvertent public disclosure and loss of professional reputation exist. Such risk is even higher when disclosure occurs in SMN which may generate voluminous word-of-mouth and facilitate rapid information transmission among key stakeholders including investors and

shareholders. In 2012, a CFO of a publicly traded company tweeted the following: "Board meeting. Good numbers=Happy Board." His Twitter account linked to other Facebook pages and Twitter accounts where ample word-of-mouth about the executive spread enormously which subsequently helped send the company stock price surging from $15 to $26.78 over the week before the earnings were released. Although this massive word-of-mouth triggered favorable reaction among investors, the board decided to terminate the contract of their CFO despite the positive effect on stock price. Carelessness in terms of public disclosure regarding financial results can place the firm and top executives in conflict with financial regulators enforcing insider trading laws. The consequences of such word-of-mouth about an executive's action and the role of SMN on this phenomenon motivates us to explore whether shareholder-oriented word-of-mouth has a stronger effect on decision and perceptions about the executives. Specifically, valuable instantaneous information release (e.g. shareholder-oriented signals) via Twitter and Facebook will influence investor and shareholder perceptions and be reflected in executives' careers such as survivability in their managerial positions. Thus;

H3a. Executive reputation established via social media networks is associated with managerial survivability such that specific shareholder-oriented reputation will increase the survivability in managerial positions.

Although we have argued that the valence of shareholder-oriented word-of-mouth is likely to send a signal to decision makers about the perception of an executive, the strength of this signal may depend on how the focal company is performing. Poor performance can alert decision makers to assess the alignment of their corporate governance (Weisbach, 1988). Similarly, when performance is good, firms will maintain the status quo and even enhance the benefits of top management team (e.g. Murphy, 1985). Thus, financial performance sends a

signal to corporate decision makers about the effectiveness of current top management team.

Financial performance signals may in turn impact the influence of executive reputation in the decision-making process. If there is a negative shareholder-oriented word-of-mouth about an executive, financial performance is one the first indicators that internal decision makers search to provide authenticity of the negative word-of-mouth. In this situation, if financial performance has been positive, it is more likely that negative word-of-mouth will be discounted and decision makers will stay inactive. Therefore, we posit that;

H3b. The effect of shareholder-oriented executive reputation in SMN on managerial survivability will be weaker when firm performance is high.

METHODOLOGY

In this section, we first illustrate the SMN setting and provide details about data resources and data collection procedures. Next, we present operationalization of the dependent and independent variables. Finally, we highlight the data analytic techniques employed to reflect the independent variables.

Social Media Networks Setting

Facebook opened its doors to all users in 2006 and has since become one of the largest social networking platforms. Facebook, considered as the largest "news" organization in the world (Gans, 2011), has over one billion active users worldwide. Although much of the interaction on Facebook occurs in private settings, word-of-mouth about products, services, and companies are mainly public and can be monitored directly. Twitter launched in 2006 as a microblogging platform hosting one of the largest online communities where the users can broadcast and consume content

(Kane et al., 2014). Twitter users broadcast and consume content by posting and reading 'tweets'. We extricate entire word-of-mouth posts about executives in Twitter since the observation platform is public by default, and permits researchers to examine multi-directional interactions among actors. This multi-platform approach allows us to examine a broad set of electronic word-of-mouth transactions and begin to understand if differences occur between platforms.

Data Collection

Dependent and Independent Variables

We build a list of all chief executive officers (CEO's) who served between July 2009 and July 2016 from the Standard & Poor's (S&P) 1500 company index using the Compustat database. We use the S&P 1500 firms because of their high visibility and large investor base, which implies that this is a suitable context for investigating word-of-mouth about highly visible executives (e.g. Hollander, Pronk & Roelofsen, 2010), especially given the importance of shareholders in our theory. Our final sample contains 125 randomly selected CEO's who held chief executive positions within this seven-year period.

For each of these executives, Crimson Hexagon was used to collect every single Twitter post and Facebook comments from July 13, 2009 until July 1, 2016 ("Crimson Hexagon" 2016). Crimson Hexagon warehouses all public Twitter and Facebook data stretching back to 2009, but removes all data that has been deleted by users. It should be noted that we collect twelve months of data for each executive based on their starting and ending date of their duty. For example, we first randomly picked a CEO name who served between 2009 and 2016. Second, if a randomly picked executive's starting day as CEO is January 2012, we picked a random serving year after January 2012 and tracked twelve month of data. Therefore, we excluded executives whose starting

date is after July 2015 and whose duty is terminated before July 2010 in order to ensure we have the full twelve months of data for each CEO, the unit of analysis in this context. We collect managerial survivability data from the Standard & Poor's Compustat database. Executive serving periods and tenure information as well as the reason for termination are included in this database. We coded survivability in managerial position as 1 if the reason for departure given in Execucomp database is "resigned".

The vast majority of studies utilizing Twitter data across all research disciplines utilize Twitter's public API to gather Twitter posts. With regards to pulling tweets from specific users, only partial data is received from Twitter's API for each user excluding data such as mention posts to other Twitter users ("GET statuses/user_timeline" 2016); this missing data could provide much needed dimensionality to word-of-mouth analysis as tweeting to a specific person versus tweeting to the general public could reveal various linguistic features important to the analysis. The Crimson Hexagon has the distinct advantage of providing full and accurate word-of-mouth scans of Twitter conversations.

Operationalization of Reputation

We follow the research stream of media communication and information systems (Aggrawal et al. 2012; Deephouse, 1996; Pollock & Rindova 2003) and use positive, neutral, and negative categories to operationalize valence of reputation about executives. We measure the valence by using LIWC (Language Inquiry Word Count) software program (Pennebaker et al. 2001), with predefined dictionaries of words to measure a variety of constructs. The LIWC dictionaries have been validated using human coders across a variety of different types of text, including online texts. In our research context, each SMN post was analyzed using the sentiment

dictionary in the LIWC program and classified as positive, negative and neutral based on the dictionary.

For the volume of word-of-mouth, we calculate the mean centered logarithm of the number of posts citing the name of executive. We use the logarithm of the number of posts to dampen the effect of extreme values.

Sub-dimensions of Reputation: Stakeholder and Shareholder Oriented Word-of-mouth

Word-of-mouth in SMN is often voluminous, unstructured, noisy, and dynamic (Gandomi & Haider, 2015). Nevertheless, SMN like Facebook and Twitter are considered valuable sources of information since people share and discuss their opinions about a certain topic freely (Medhat, Hassan, & Korashy, 2014). Despite the fact that most online users are regarded as passive readers, opinions and word-of-mouth have been shown to influence each other (Goes et al. 2014). Therefore, public comments about executives and companies posted in SMN require deeper analysis of quality (Lau, Li, & Liao, 2014). In this study, we seek a comprehensive understanding of what public opinions actually are that form the overall executive reputation. Specifically, instead of relying solely on the overall valence of reputation, we intend to tease out what type of content actually drives the major reputation valence. Considering the dual roles of executives –spokesman of a company and social influencer, several different groups of audiences may talk and provide opinion about executives including customers, journalists, shareholders, peer directors, fans, prospective employees, and social communities. We argue that distinct audience groups will affect the spread of word-of-mouth and influence the firm related outcomes in different ways. In other words, stakeholder oriented word-of-mouth may differ from shareholder oriented word-of-mouth on firm practices. For instance, customers may not generate word-of-mouth about executive

performance related to earning calls yet may value and contribute more if word-of-mouth is about an executive's social responsibility. Thus, we need a robust approach to observe and quantify such diverse signals and incorporate it into our methodology.

Our methodology intended to measure the association of public reputation of executives with managerial level outcomes. We follow information systems research stream (e.g., Miranda, Young, & Yetgin, 2016) and applied methodologies (e.g., Aggarwal et al. 2012) to operationalize hypothesis 1 by drawing on mainstream text mining tools. Testing second and last hypotheses require deeper understanding of the effects of each separate constituent of reputation construct. Thus, we applied a relatively novel data mining methodology, called feature-based opinion mining (Eirinaki, Pisal, & Singh, 2012) to extract and quantify explanatory variables. To extract stakeholder and shareholder oriented opinions, we apply opinion mining, ranking, and classifying algorithms for SMN data in three steps. The first algorithm identifies the most important features of word-of-mouth about an executive, the second one ranks the valence of each feature, and the last one classifies the output into stakeholder and shareholder oriented word-of-mouth sub-dimensions. Since multi-faceted conversations have hundreds of opinions and a lot of noise, we argue that some features are more important than the others and these features distinguish one dimension from other less important dimensions. Thus, instead of solely relying on overall valence of word-of-mouth, we uncover important sub-dimensions with their own valence rankings. Then we classify these important topics into stakeholder and shareholder oriented coverage dimensions. Because we argue that the effect of word-of-mouth will vary for each sub-dimensions, it is essential to tease out which sub-dimension has different valence. Our assumption is that online users frequently comment on important features of an executive and these specific features may have a greater influence on executive reputation.

Data Processing

We use feature-based opinion mining model (Eirinaki et al. 2012) that will extract useful information related to the word-of-mouth about executives and classify it as either stakeholder oriented or shareholder oriented signals along with their ranked sentiments. We depict the main elements of our data mining approach in Figure 2.2.

```
┌──────────┐      ┌──────────────┐
│ Data Set │─────▶│     Data     │
└──────────┘      │ Pre-Processing│
                  └──────────────┘
                          │
                          ▼
                  ┌──────────────┐      ┌──────────────┐
                  │Opinion Mining│◀────▶│   Opinion    │
                  │    Engine    │      │   Ranking    │
                  └──────────────┘      │  Algorithms  │
                          │             └──────────────┘
                          ▼
                  ┌──────────────┐
                  │Classification of│
                  │    Ranked    │
                  │   Opinions   │
                  └──────────────┘
```
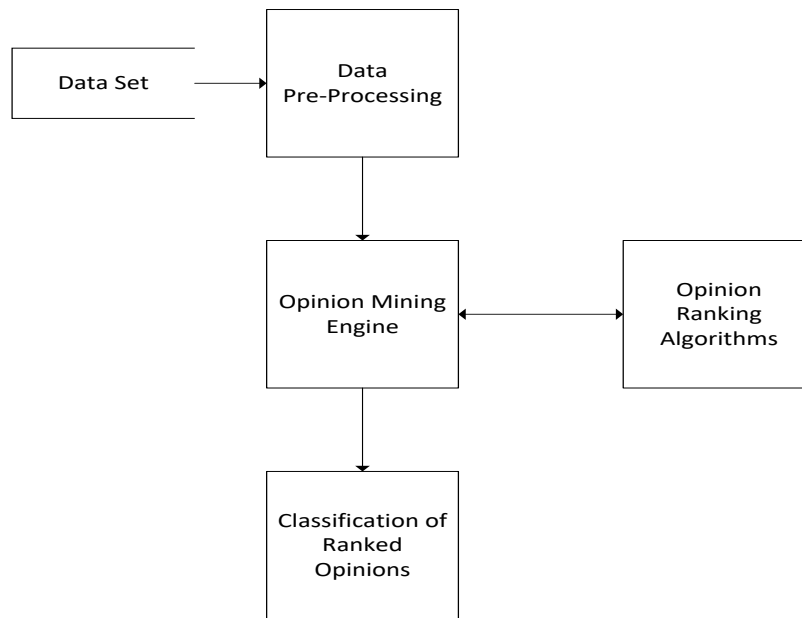
Figure 2.2. Model Architecture (Eirinaki et al. 2012)

Data Preprocessing: We clean up the raw data set by applying pre-processing to remove stop words, stemming, and punctuation and transform it to a computational format by using natural language toolkit (NLTK) Python programming language (Bird, Loper, & Klein, 2009).

Opinion Mining Engine: The opinion mining engine includes a POS (parts-of-speech) tagger and word tokenize modules used to process the text, such as, marking up a word in a conversation as

corresponding to part of speech, and computing the distance between a noun and its closest adjective.

Opinion Ranking Algorithms: We operate ranking algorithms that rank the users' opinions based on the scores attributed to the extracted features. These scores are used to determine the orientation of the word-of-mouth. The details of the ranking algorithms will be presented as a separate section below.

Classification of Ranked Opinions: We classify the final ranked output into shareholder and stakeholder sub-groups. We compiled a dictionary, provided in the supporting information section, and utilized word roots to identify group matches from the final output.

Opinion Mining and Ranking Algorithms

High Adjective Count Algorithm

For our model, a feature-based implementation, in other words, an algorithm for the identification of the most relevant features is necessary. These features are mainly represented by nouns, and the dominant sentiment is conveyed by an adjective (Hu & Liu, 2004). For the feature selection task, we draw on Eirinaki et al. (2012) who identified potential features with the high adjective count (HAC) algorithm. Pseudo code of the HAC algorithm provided in Figure 2.3.

82

```
map_noun_scores = { }
feature list = []
for posts in corpus:
    apply stemming to posts
    apply POS tagger to posts
    for sentences in posts:
        if sentences have nouns and adjectives
        find the closest noun for each adjective in terms (argmin dist(adjective, noun))
        map_noun_scores++
for nouns in map_noun_scores:
    if noun_score > threshold
    append that noun to feature list
```

Figure 2.3. High Adjective Count Algorithm

The core idea of the algorithm is that the featured nouns for which users talk and share many opinions are most likely to be the important and distinguishing features than those for which users do not express such opinions (Eirinaki et al. 2012). Posts refer to the posts users share to form word-of-mouth on Twitter and Facebook. Sentences refer the sentence of a post. We first apply Porter (1980) stemming and use pre-trained POS tagger to determine the part of speech for each word. Each adjective is associated with the noun to which it is the closest. That adjective will most likely define this noun. We increase the score of the noun by one for each such adjective. After iterating through the whole text corpus, the algorithm will have allocated scores for each of the nouns. Subsequently, we refer to these as opinion scores when the score is greater than a predetermined threshold (Eirinaki et al. 2012).

The Max Opinion Score Algorithm

Next step is identifying the valence of extracted features. For this, we follow the Maximum

Opinion Score (MOS) algorithm proposed by Eirinaki et al. (2012) as in Figure 2.4.

```
potential_features = { }
opinion_words = { }
inversions = { }
positive_score_feature = 0
negative_score_feature = 0
    for posts in corpus:
    apply stemming to posts
    for sentences in posts:
            if sentences have any potential feature (mark as F) and opinion_words (mark as O)
      for each opinion word
            find the closest potential feature for each opinion word (argmin dist (O, F))
            if any inversion word is located in the left context:
                    if opinion word has a positive sentiment: negative_score_feature ++
                    else: positive_score_feature ++
            else
                    if opinion word has a positive sentiment: positive_score_feature ++
                    else: positive_score_feature ++
```

Figure 2.4. The Maximum Opinion Score Algorithm

We label the sentiment defining adjectives as opinion words. In this study, a commonly

known source Senti-WordNet serves as the opinion words dictionary (Baccianella et al. 2010).

Since the adjectives split into positive and negative valence, inversion words like "not" in the

context of these adjectives reverse the meaning of the word valence. Altogether, we aim to extract

the context which includes at least an opinion word and a feature derived by HAC algorithm. The

position of opinion word is detected and the score of the closest feature is computed in accordance

with the valence of the opinion word. Finally, we aggregated the resulting positive and negative

scores and provided a weighted net score with n as number of features as follows where X refers to positive feature score and Y refers to negative feature score:

$$\text{Net Score} = \frac{\sum_{i=1}^{n} X_i - y_i}{\sum_{i=1}^{n} X_i + y_i}$$

Net score is scaled on the interval from -1 (all the word-of-mouth to that feature is negative) to +1 (all the word-of-mouth to that feature is positive). We then classified relevant features along with their aggregated scores into *shareholder* and *stakeholder* oriented groups and quantified the values within each group.

Moderating Variables

Volume of word-of-mouth: We measure the volume of word-of-mouth with the mean centered logarithm of the number of posts citing the name of an executive. A higher number of posts may attract more board attention and enhanced knowledge about the public reputation of the executive. We use the logarithm of the number of posts to reduce the effect of extreme values.

Financial performance: We use stock response modeling, which provides evidence pertaining to stock returns as a measure of financial value (Schneiderjans et al. 2013). We derive earnings per share (EPS) difference between the start and end of year values of the observation period to reflect the financial impact of executive word-of-mouth.

Control Variables

A number of control variables are included in the analyses. Tenure in executive position has been shown to significantly affect the decisions of corporate leaders (Hambrick & Mason, 1984). We use the number of years that a current CEO had been in the office as our tenure measure. In addition, we use demographic variables, such as age and gender, to account for individual

differences (e.g. Lee & James, 2007). We also consider the fact that executives who founded their firms might have different rights in their firms than non-founders. We obtain this information from SEC filings of the firm and use this information in our analyses as a categorical variable. Finally, we control for an external underlying scandal as a binary variable that could be driving the word-of-mouth. We generate a categorical variable that counts if the word-of-mouth specifically contains the word "scandal" (Bednar, 2012).

EMPIRICAL ANALYSIS

Our empirical approach is to employ logistic regression analysis to estimate the effects of the word-of-mouth in two separate contexts; first, we examine the overall effect of valence of word-of-mouth on survivability in executive position. Second, we tease out the specific orientation of the coverage by splitting overall word-of-mouth into three dimensions, stakeholder-oriented, shareholder-oriented and the remaining irrelevant coverage (coded as *others*). We run the first model by using only overall valence and control variables. We run the second model by using the aforementioned three dimensions along with the interactions in order to monitor the specific effects of each dimension that constitutes overall valence. Descriptive statistics for our measures are provided in Table 2.1 below. To examine whether explanatory variables are likely to cause collinearity concerns, Spearman rank correlations were computed for these measures. Scandal variable is omitted due to multicollinearity issue and lack of sufficient variance resulting in 122 observations in the sample. These correlations are shown in Table 2.2. All correlations are less than 0.5, which indicates that multicollinearity across the IM tactics is less likely (Kishore, Agrawal, & Rao, 2004). The extremity range of valence is from -100 to +100 simply states the percentage strength of the coverage based on its positive or negative sign.

Table 2.1: Descriptive Statistics (n=125)

	mean	sd	min	max
Survivability	.459	.5	0	1
Volume (log)	3.15	.939	1.2	4.6
Shareholder Valence	.68	5.5	-16	18
Stakeholder Valence	.55	13.59	-51	72
Valence Overall	.565	16.2	-45	77
Age	64.48	6.15	49	85
Tenure	10.07	7.9	1	40
Financial Performance	4.543	10.58	-35.8	53.8
Founder	.204	.405	0	1
Gender	.09	.24	0	1
Scandal	.024	.15	0	1

Table 2.2: Spearman's Correlations for Final Model

Survivability	1.00						
Shareholder	0.47	1.00					
Stakeholder	0.35	0.33	1.00				
Financial Performance	0.29	0.4	0.18	1.00			
Founder	0.38	0.04	0.19	0.11	1.00		
Tenure	-0.13	0.1	0.01	0.05	-0.1	1.00	
Age	-0.3	-0.06	-0.24	0.04	-0.16	0.3	1.00

As a second check for multi-collinearity problems, we calculate variance inflation factors (VIF) for each variable. The average VIF values are displayed in Table 2.3 which are less than the acceptability threshold of 5 (Hair et al., 2006), implying that multicollinearity is not a concern. Further, to limit potential concerns about unequal variances of our explanatory variables and controls across the range of the survivability measure, we employ heteroskedasticity-consistent standard errors (White, 1980). Finally, our choice of a lagged-model, wherein word-of-mouth and moderating variables are measured in a period prior to the period for the dependent variable measurement, should limit concerns of endogeneity.

Table 2.3: Variance Inflation Factors (n=122)

Variable	VIF	Sqrt VIF	Toler.	R-Squared
Survivability	1.83	1.35	0.5458	0.4542
Stakeholder	1.39	1.18	0.7175	0.2825
Shareholder	1.49	1.22	0.6731	0.3269
Volumelog	1.15	1.07	0.8714	0.1286
Age	1.46	1.21	0.6864	0.3136
Tenure	1.46	1.21	0.6863	0.3137
Fin. Per.	1.19	1.09	0.8408	0.1592
Founder	1.25	1.12	0.7996	0.2004
Mean VIF	1.40			

Results of Logistic Regression Estimations

Using logistic regression analysis, we estimate the effects on survivability in executive position of word-of-mouth in three step-wise models; we first examine the effect of overall valence of word-of-mouth on survivability. Second, we tease out the specific orientation of the coverage by splitting overall word-of-mouth into three dimensions, stakeholder-oriented, shareholder-oriented and the remaining irrelevant coverage (coded as *others*). Finally, we include the interactions of volume and financial performance in the third model. Results of the regression analysis are provided in Table 2.4.

We apply robust standard errors (Rogers, 1994) and find that the overall model is significant with a chi-square value of 28.54 significant at $p < 0.001$. The model fit is 56%. Model (1) only includes the overall valence and control variables. Overall valence of word-of-mouth about the executives has a positive, statistically significant effect on their survivability in current positions. For instance, for every additional score in overall valence, we expect a 0.033 increase in the log odds of survivability in the managerial position with $p < 0.05$.

Model (2) introduces sub-dimensions of overall valence. We see a statistically significant relation between valence of stakeholder-oriented word-of-mouth and survivability with $p<0.05$ and valence of shareholder-oriented word-of-mouth with $p<0.01$. We also observe the effect of noisy coverage, neither shareholder nor stakeholder oriented valence, and find that irrelevant word-of-mouth does not have a statistically significant impact on survivability. In order to understand if omitted interaction effects may be the driver of these significant results, we ran Model (3) by incorporating the moderation variables.

In Model (3), we first include and interaction between shareholder-oriented coverage and financial performance. Although we observe a negative moderating effect as we hypothesized associated with the impact of valence of shareholder-oriented coverage will be higher when financial performance is poor, it is not statistically significant. Second, we see a statistically significant positive effect of volume moderator on survivability within the stakeholder-oriented context. As shown in Model (3), we note that stakeholder-oriented valence is positively and significantly associated with the managerial survivability when the volume of word-of-mouth is high. We suspect that most the variance of Stakeholder-oriented coverage in Model (2) is absorbed by the volume interaction. However, shareholder-oriented valence is positively and significantly associated with the managerial survivability regardless of the financial situation of the firm.

89

Table 2.4: Logistic Regression Estimates of Survivability (n=122)

	(1) Survivability	(2) Survivability	(3) Survivability
Overall Valence	0.0329*		
	(2.00)		
Volume (log)	-0.123	0.0799	-0.0772
	(-0.49)	(0.27)	(-0.23)
Age	-0.104*	-0.0828	-0.0701
	(-2.16)	(-1.47)	(-1.23)
Tenure	0.0252	-0.0261	-0.0314
	(0.69)	(-0.46)	(-0.59)
Financial Per.	0.0703*	0.0617*	0.0697*
	(2.47)	(2.45)	(2.32)
Founder	1.948**	2.328**	2.838**
	(3.04)	(3.27)	(3.02)
Stakeholder Valence		0.173*	0.257
		(2.54)	(1.34)
Shareholder Valence		0.293**	0.321**
		(3.25)	(2.69)
Other Valence		-0.0781	-0.0833
		(-1.77)	(-1.93)
Stakeholder*Volume			2.233*
			(2.21)
Shareholder*Fin. Perf.			-0.00217
			(-0.23)
_cons	5.954	3.943	3.929
	(1.80)	(1.10)	(1.06)
N	122	122	122

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

DISCUSSION

The goal of this study is to examine the effect of executives' public reputation in online networking platforms as an external cue on managerial survivability. Based on our theory and empirical findings, we observed initial evidence that the valence of word-of-mouth about an executive in online networking platforms such as Facebook and Twitter is a salient predictor of managerial survivability. We also investigated the specific dimensions of public word-of-mouth in order to understand the main drivers of the valence of overall word-of-mouth and found that stakeholder and shareholder-oriented coverage have significant effects under some circumstances. On one hand, stakeholder-oriented coverage is not statistically significant in the final model, yet it is positively and statistically associated with managerial survivability when the volume of coverage is high. On the other hand, shareholder-oriented coverage maintains its significance even after introducing the financial performance interaction. Finally, we observed that financial performance has no statistically significant effect as a moderator. Table 2.5 summarizes our results.

Table 2.5: Summary of Hypotheses Tests

Variable	Predicted Effect	Result of Hypothesis Testing
Overall Valence of Reputation	Positive	Supported
Stakeholder Valence	Positive	Unsupported
Shareholder Valence	Positive	Supported
Moderator (Volume)	Positive	Supported
Moderator (Financial Performance)	Negative	Unsupported

Our findings indicate that one of the main drivers of overall reputation's effect on survivability is shareholder-oriented word-of-mouth. As we posit in hypothesis 3, when people talk and post about the performance and managerial capability of an executive in a shareholder related context,

internal decision makers pay more attention to the valence coverage. Our view is that, regardless of financial performance influence, word-of-mouth, which is generated and spread by financial interest groups such as analysts, investors, and business followers, plays an important role on career decisions about executives. Poor or high financial performance is not affecting the strength of this relation. A reasonable explanation for this finding can be alternative positive deeds of an executive. For example, a firm, which makes considerable investment on R&D projects, may not recognize a return on investment in the short-term. However, people may comment and post positive things about the executive leadership anticipating stronger performance later. Alternatively, people can share negative thoughts about the performance of an executive although earnings per shares increases during the observation window.

Our second important finding is that stakeholder-oriented word-of-mouth is significant when volume of the word-of-mouth is high. This finding suggests that, dissident stakeholders that collectively use SMN can increase attention to their causes and prompt firm decision makers response to their concerns. This finding also reveals that the stakeholder-oriented word-of-mouth may not be considered as a driver of overall valence of reputation if there is not enough coverage in terms of volume.

Finally, our results indicate that the overall valence of reputation in SMN can be influential for the executives' career. Our results corroborate the previous findings in conventional media that media coverage plays a significant role on CEO dismissal or survival (Bednar, 2012), but extend these findings to show pertinence in social media networks.

CONCLUSION

CEO and company reputation are inseparably linked and corporate reputation may not be insulated from the public reputation of a company's top leadership. Consequently, CEO reputation impacts not only a firm's value (Francis et al. 2008) but also to the managers' professional career outcomes specifically when media serves as a considerable corporate governance mechanism (Bednar, 2012). For that reason, companies should therefore be concerned with corporate reputation for the long-term as well as the public reputation of top executives to maximize competitive advantage. Likewise, top executives should also be concerned with their own public reputation to increase their career outcomes.

In this study, we reveal the influence of public reputation of top executives in social media networks. We find that there is an association between the valence of reputation and managerial survivability. As the main limitation of this study, we acknowledge that there are many other internal and external reasons that may affect managerial decision making. However, approaching this phenomenon from a specific perspective, we focus on highlighting the effect of public reputation enabled by recent advances in social information technology platforms. We also limited ourselves and made an effort to analyze reputation within two contexts; future research may reveal other dimensions of media coverage within the executives' realm. Finally, we directly used a key word dictionary in order to categorize the coverage context. Future research may use recently developed word-embeddings deep learning techniques to enhance the comprehensiveness of the intended categorization. Nonetheless, our methodological design provides initial evidence about the potential implications of word-of-mouth by different groups.

The main implications of the study and contribution to the literature are twofold. First, there is no known prior research which has explored the reputation and executive survivability

relationship within SMN domain. SMN includes a wide spectrum of information aggregated from various audiences ranging from investment analysts, investors, shareholders, social communities, agencies, employees to current and prospective customers. Together, SMN plays an extended social arbiter role and is considered as a credible and prominent platform for rendering public assessments of firms and the individuals associated with them (Aggarwal et al. 2012). Therefore, this study provides insights into the external drivers of top management team alignments by investigating executive reputation unexplored in a social media setting previously.

Second, our study explored the executive reputation construct in specific contexts, namely stakeholder and shareholder orientation, to identify the main the features of the overall reputation by using novel data analytic techniques. We believe our methodology is novel and timely for the following reasons. First, we complement prior research in this stream that uses self-reported survey data and predefined computer-aided tools (e.g., citation…). A survey data approach may be subject to data limitations pertaining to sample size, recall biases, and low response rates (Bolino et al., 2008) and predefined computer-aided tools might not be capable of extracting targeted insights located in unstructured datasets. Moreover, since public's views of an executive may be originated from any user group, it is possible to capture the buried opinions of broad audiences by analyzing social network communication.

REFERENCES

Aggarwal, R., Gopal, R., Gupta, A., and Singh, H. 2012. "Putting Money Where the Mouths Are: The Relation between Venture Financing and Electronic Word-of-Mouth", *Information Systems Research,* (23:3), pp. 976-992.

Aggarwal, R., and Singh, H. 2013. "Differential Influence of Blogs Across Different Stages of Decision Making: The Case of Venture Capitalists," *MIS Quarterly,* (37:4), pp. 1093-1112.

Agrawal, A., and Knoeber, C. R. 2001. "Do Some Outside Directors Play a Political Role?" *Journal of Law and Economics,* (44), pp. 179-198.

Arvidsson, A., 2006. "Brand value" *Journal of Brand management*, (13:3), pp.188-192.

Baccianella, S., Esuli, A. and Sebastiani, F., 2010, May. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. *In LREC*, 10, pp. 2200-2204.

Bansal, H.S. and Voyer, P.A., 2000. "Word-of-mouth processes within a services purchase decision context" *Journal of Service Research*, (3:2), pp.166-177.

Berger, K., Klier, J., Klier, M., & Probst, F. 2014. "A Review of Information Systems Research on Online Social Networks" *Communications of the Association for Information Systems*, (35:1), pp. 8-14.

Bird, Steven, Edward Loper and Ewan Klein. 2009. Natural Language Processing with Python. *O'Reilly Media Inc*.

Bollen, J., Mao, H., & Zeng, X. 2011. "Twitter Mood Predicts the Stock Market" *Journal of Computational Science*, (2:1), pp. 1-8.

Boyd, D. and Ellison, N., 2010. "Social network sites: definition, history, and scholarship" *IEEE Engineering Management Review*, (3:38), pp.16-31.

Burt, R. S. 2004. "Structural holes and good ideas*" American Journal of Sociology*, (110:2), pp. 349-399.

Core, J. E., Guay, W., & Larcker, D. F. 2008. "The power of the pen and executive compensation" *Journal of Financial Economics*, (88:1), pp. 1-25.

Deephouse, D. L. 2000. "Media reputation as a strategic resource: An integration of mass communication and resource-based theories" *Journal of Management*, (26:6), pp. 1091-1112.

Dellarocas, C., 2003. "The digitization of word of mouth: Promise and challenges of online feedback mechanisms" *Management Science,* (49:10), pp.1407-1424.

Dhaliwal, D.S., Radhakrishnan, S., Tsang, A. and Yang, Y.G., 2012. "Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure" *The Accounting Review,* (87:3), pp.723-759.

Donaldson, T. and Preston, L.E., 1995. "The stakeholder theory of the corporation: Concepts, evidence, and implications" *Academy of Management Review,* (20:1), pp.65-91.

Eirinaki, M., Pisal, S. and Singh, J., 2012. "Feature-based opinion mining and ranking" *Journal of Computer and System Sciences*, (78:4), pp.1175-1184.

Ellison, N. B., and boyd, d. 2013. Sociality through Social Network Sites. *Oxford University Press*, pp. 151-172.

Evans, D., 2012. Social media marketing: An hour a day. *John Wiley & Sons*.

Gaines-Ross, L., 2000. "CEO reputation: A key factor in shareholder value" *Corporate Reputation Review*, (3:4), pp.366-370.

Gandomi, A. and Haider, M., 2015. "Beyond the hype: Big data concepts, methods, and analytics" *International Journal of Information Management*, (35:2), pp.137-144

Geletkanycz, M. A., Boyd, B. K., & Finkelstein, S. 2001. "The strategic value of CEO external

directorate networks: Implications for CEO compensation" *Strategic Management Journal*, (22:9), pp. 889-898.

Gilly, M. C., Graham, J. L., Wolfinbarger, M. F., & Yale, L. J. 1998. "A dyadic study of interpersonal information search" *Journal of the Academy of Marketing Science*, 26, pp. 83-10

Gitlin, T., 1980. "The whole world is watching: Mass media in the making & unmaking of the new left" *University of California Press*.

Goes, P.B., Lin, M. and Au Yeung, C.M., 2014. "Popularity Effect in User-Generated Content: Evidence from Online Product Reviews" *Information Systems Research*, (25:2), pp. 222-238.

Fiske, S.T. and Taylor, S.E., 1991. Social cognition, *2nd. NY: McGraw-Hill,* pp.16-15.

Fombrun, C., & Shanley, M. 1990. "What's in a name? Reputation building and corporate strategy" *Academy of Management Journal,* (33:2), pp. 233-258.

Fombrun, C., 1996. Reputation. *John Wiley & Sons, Ltd.*

Fombrun, C. J., Gardberg, N. A., & Sever, J. M. 2000. "The reputation quotient: A multi-stakeholder measure of corporate reputation" *Journal of Brand Management*, (7:4), pp. 241-255.

Francis, J., Huang, A.H., Rajgopal, S. and Zang, A.Y., 2008. "CEO Reputation and Earnings Quality". *Contemporary Accounting Research,* (25:1), pp. 109-147.

Fredrickson, J.W., Hambrick, D.C. and Baumrin, S., 1988. "A model of CEO Dismissal" *Academy of Management Review*, (13:2), pp. 255-270.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., and Tatham, RL. 2006. *Multivariate Data Analysis*, Pearson Prentice Hall.

Hambrick, D.C. and Mason, P.A., 1984. "Upper echelons: The organization as a reflection of its top managers" *Academy of Management Review*, (9:2), pp. 193-206.

Heath, C. 1996. "Do people prefer to pass along good or bad news? Valence and relevance of news as predictors of transmission propensity" *Organizational Behavior and Human Decision Processes*, 68, pp. 79-94.

Herr, P.M., Kardes, F.R. and Kim, J., 1991. "Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnosticity perspective" *Journal of Consumer Research*, (17:4), pp. 454-462.

Hall, R., 1992. "The strategic analysis of intangible resources" *Strategic Management Journal*, (13:2), pp. 135-144.

Hayward, M. L., Rindova, V. P., & Pollock, T. G. 2004. "Believing one's own press: The causes and consequences of CEO celebrity" *Strategic Management Journal*, (25:7), pp. 637-653.

Hu, M. and Liu, B., 2004. "Mining and summarizing customer reviews" *In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 168-177. ACM.

Jansen, B.J., Zhang, M., Sobel, K. and Chowdury, A., 2009. "Twitter power: Tweets as electronic word of mouth" *Journal of the American Society for Information Science and Technology*, (60:11), pp. 2169-2188.

Kane, G. C., Alavi, M., Labianca, G. J., and Borgatti, S. 2014. "What's Different about Social Media Networks? A Framework and Research Agenda," *MIS Quarterly* (38:1), pp. 274-304.

Kosicki, G.M., 1993. "Problems and opportunities in agenda-setting research" *Journal of Communication*, (43:2), pp.100-127.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. 2011. "Social Media? Get serious! Understanding the functional building blocks of social media" *Business Horizons*, (54:3), pp. 241-251.

Kietzmann, J. H., Silvestre, B. S., McCarthy, I. P., & Pitt, L. F. 2012. "Unpacking the social media phenomenon: towards a research agenda" *Journal of Public Affairs*, (12:2), pp. 109-119.

Kirchmaier, T., & Stathopoulos, K. 2008. "From fiction to fact: The impact of CEO social networks" *Manchester Business School working paper*. 537

Kushin, M. J., & Yamamoto, M. 2010. "Did social media really matter? College students' use of online media and political decision making in the 2008 election" *Mass Communication and Society*, (13:5), pp. 608-630.

Lau, R.Y., Li, C. and Liao, S.S., 2014. "Social analytics: Learning fuzzy product ontologies for aspect-oriented sentiment analysis" *Decision Support Systems*, 65, pp. 80-94.

Lee, P.M. and James, E.H., 2007. "She'-e-os: Gender Effects and Investor Reactions to the Announcements of Top Executive Appointments," *Strategic Management Journal* (28:3), pp. 227-241.

Lin, N. 1999. "Building a network theory of social capital" *Connections*, (22:1), pp. 28-51.

Marwick, A.E., 2013. Status update: Celebrity, publicity, and branding in the social media age. *Yale University Press.*

Medhat, W., Hassan, A. and Korashy, H., 2014. "Sentiment analysis algorithms and applications: A survey" *Ain Shams Engineering Journal*, (5:4), pp. 1093-1113.

McCombs, M., Llamas, J. P., Lopez-Escobar, E., and Rey, F. 1997. "Candidate Images in Special Elections: Second Level Agenda Setting Effects" *Journalism and Mass Communication Quarterly* (74), pp.703–717.

Mehra, A., Dixon, A.L., Brass, D.J. and Robertson, B., 2006. "The social network ties of group leaders: Implications for group performance and leader reputation" *Organization Science*, (17:1), pp. 64-79.

Milbourn, T. T. (2003). "CEO reputation and stock-based compensation" *Journal of Financial Economics*, (68:2), pp. 233-262.

Mills, C. W. 1959. The Sociological Imagination, *Oxford University Press*, London, UK.

Miranda, S.M., Young, A. and Yetgin, E., 2016. "Are social media emancipatory or hegemonic? Societal effects of mass media digitization" *MIS Quarterly*, (40:2), pp. 303-329.

Mishina, Y., Block, E.S. and Mannor, M.J., 2012. "The path dependence of organizational reputation: How social judgment influences assessments of capability and character" *Strategic Management Journal*, (33:5), pp. 459-477.

Murphy, Kevin J. 1985 "Corporate performance and managerial remuneration: An empirical analysis" *Journal of Accounting and Economics,* (7-1), pp. 11-42.

Oh, O., Agrawal, M., & Rao, H. R. 2013. "Community intelligence and social media services: a rumor theoretic analysis of tweets during social crises" *MIS Quarterly*, (37:2), pp. 407-426.

Rogers, W., 1994. "Regression Standard Errors in Clustered Samples," *Stata Technical Bulletin* (3:13).

Pennebaker, J.W., Francis, M.E. and Booth, R.J., 2001. Linguistic inquiry and word count: LIWC 2001. *Mahway: Lawrence Erlbaum Associates,* 71, p. 2001.

Penrose, E. T. 1959. The Theory of the Growth of the Firm. *New York: John Wiley*.

Podolny, J. M. 2001. "Networks as the Pipes and Prisms of the Market" *American Journal of Sociology,* (107:1), pp. 33-60.

Pollock, T. G., and Rindova, V. P. 2003. "Media Legitimation Effects in the Market for Initial Public Offerings," *Academy of Management Journal* (46:5), pp. 631-642.

Porter, M.F., 1980. "An algorithm for suffix stripping" *Program,* (14:3), pp. 130-137.

Rein, I., Kotler, P., Hamlin, M. and Stoller, M., 2006. High visibility: Transforming your personal

97

and professional brand. *McGraw Hill Professional.*

Sedereviciute, K., & Valentini, C. 2011. "Towards a More Holistic Stakeholder Analysis Approach. Mapping Known and Undiscovered Stakeholders from Social Media" *International Journal of Strategic Communication,* (5:4), pp. 221-239.

Spence, M., 1973. "Job market signaling" *The Quarterly Journal of Economics,* pp.355-374.

Tufekci, Z., & Wilson, C. 2012. "Social media and the decision to participate in political protest: Observations from Tahrir Square" *Journal of Communication,* (62:2), pp. 363-379.

Van Hoye, G. and Lievens, F., 2009. "Tapping the grapevine: a closer look at word-of-mouth as a recruitment source" *Journal of Applied Psychology,* (94:2), p.341.

Wade, J. B., Porac, J. F., Pollock, T. G., & Graffin, S. D. 2006. "The Burden of Celebrity: The Impact of CEO Certification Contests on CEO Pay and Performance" *Academy of Management Journal,* (49:4), pp. 643-660.

Wernerfelt, B. 1984. ″A Resource‐Based View of the Firm″ *Strategic Management Journal*, (5:2), pp. 171-180.

Westphal, J. D., and Deephouse. D. L. 2011. "Avoiding Bad Press: Interpersonal Influence in Relations between CEOS and Journalists and The Consequences for Press Reporting about Firms and Their Leadership," *Organization Science*, (22:4), pp. 1061-1086.

Westphal, J. D., Park, S. H., McDonald, M. L., & Hayward, M. L. 2012. "Helping other CEOs Avoid Bad Press Social Exchange and Impression Management Support among CEOs in Communications with Journalists" *Administrative Science Quarterly,* (57:2), pp. 217-268.

White, H., 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," *Econometrica* (48:4) pp.817-838.

Zajonc, R.B., 1968. "Attitudinal effects of mere exposure" *Journal of Personality and Social Psychology*, (9:2p2), p.1.

SUPPORTING INFORMATION

Stakeholder and Shareholder Key Word Dictionary

STAKEHOLDER				SHAREHOLDER		
Broad Stakeholder	**Community**	**Employees**	**Environment**	acqui*	grow*	securities
education	charit*	agents	biodiversity	asset*	income	shareh*
justice	communit*	associates	clean	bankers	industry	shareo*
purpos*	farmer*	colleagues	climate change	board	infrastructure	stock*
responsibility	improve lives	diversity	energy	business*	invest*	strateg*
responsible	philanthrop*	employee*	environment	buyback*	liquidity	tax
stakeholder	small business	managers	environmenta*	capabilities	loan*	trading
veterans	societ*	people	preservation	capital*	managem*	transaction*
accountability	**Customer**	safety	renewable	cash*	margins	gain*
africa	client*	team*	sustainabil*	competi*	market*	saving*
america*	consumer*	voluntee*	sustainable	cost*	money	
crisis	customer*	wom*	water	credit*	mortgage	
culture	patients	work*	wind	currency	operatio*	
director*	product	**Government**	global warming	debt	operating margins	
diversity	safe*	compliance		deposit*	opportunities	
education	price	governance		diversified	owner*	
europe	discount	government*		dividend*	payment*	
families	promotion	regulati*		earning*	performance	
gain		regulato*		econom*	portfolio	
global*				effective*	product*	
partner*				efficien*	profit*	
partnership				equity	projects	
recession				expend*	repurchase*	
relationships				expense*	results	
value				finance*	return	
ethics				financial*	revenue*	
school				fiscal	risk	
health*				fund*	sale*	

Chapter 3.

A BRAND NEW LOOK AT YOU: PREDICTING BRAND PERSONALITY IN SOCIAL MEDIA NETWORKS WITH MACHINE LEARNING

ABSTRACT

Compared with the wealth of research focused on automated human personality assessment, surprisingly little research has focused on advancing methods for obtaining brand personality from social media content. Brand personality is a nuanced aspect of the brand that has a consistent set of traits aside from its functional benefits. In this study, we introduce a novel, automated and highly generalizable data analytics approach to extract near real-time estimates of brand personalities in social media networks. Our new approach uses a hybrid machine learning algorithmic design, which bypasses often extensive manual coding tasks, thus providing an adaptable and scalable tool that can be used for a range of management studies. Our proposed approach may have strong implications for academic scholars in enhancing the theoretical understanding of channeled and perceived brand personality embodied in social media networks. Moreover, we aim to provide additional benefits to practitioners including the ability to foster branding strategies by utilizing big data resources. To the best of our knowledge, there are no existing machine learning approaches developed for brand personality prediction.

Keywords: brand personality, social media networks, data analytics, machine learning

INTRODUCTION

Brand names are regarded among the most valuable assets owned by a firm. Strong and distinguished brands significantly enhance firm performance (Madden, Fehle, & Fournier, 2006). For every brand, the relationship with their users is key. In fact, it has been shown that brands can exhibit personalities, "brand personality", similar to how human beings exhibit personalities (Aaker, 1997). As brands build these personalities, people have actually been shown to interact with brands as if they were human (Levy, 1985). Naturally, consumers seek brands with personalities that are congruent with either their own or their aspirational (ideal) personalities (Sirgy, 1982).

The growing plethora of social media networks (SMN) have sparked an opportunity to understand how firms foster their branding efforts. Due to the growing potential for SMN to be utilized as efficient marketing and brand-building platforms, firms have increasingly been moving their branding efforts to this digital interactive medium. As a result, social branding has become an essential form of marketing communication to convey core brand personality. Having the ability to use an effective marketing communications strategy to distinguish itself from competitors has become a requisite to enhance customer relationship and foster brand equity. Thus, for a firm to understand their brand personality on SMN, they must have some capability to assess the channeled and perceived brand personality through their content generation and interactive dialog with their consumers.

Since 1997, most of the marketing literature has embraced self-reporting tools (e.g., Likert scale surveys) based on Aaker's scale to assess brand personality. Such self-reporting tools are often expensive, labor-intensive and time-consuming. They exhibit bias issues, and the results can become outdated very quickly. In this age of data driven analytics, brand personalities are also

being projected real-time on brand's social media accounts, and traditional methods of surveying brand personality cannot cope with the speed of brand social media content creation.

In general, research in personality on SMN is positioned at the intersection of individuals, organizations, and technology. Thus, using advanced analytics to understand social data is an emerging research field across different academic disciplines including psychology, marketing, management, and information systems. As a result, personality research, both at the individual and organizational level, has become a widely studied topic and several analytic methods have been developed to assess personality in various contexts.

Typically, extant analytic methods require extensive content customization and static closed vocabulary approaches show limitations in terms of comprehensiveness. Some recent works (e.g. Park et al. 2015) have conducted automated personality assessments by using open vocabulary approaches -integrating unsupervised machine learning techniques with multiple feature selection methods- to build robust language models in SMN. Despite rigorous research efforts in human personality assessment in social media content, studies are limited in the brand personality domain. Thus, we were motivated to develop a data analytics approach to detect and analyze social psychological constructs, such as brand personality, from SMN content.

Data analytic implementations are relatively rare in marketing literature (with notable exceptions including Culotta & Cutler, 2016; Huang & Loa, 2016; Jacobs, Donkers, & Fok, 2016) and there are no extant approaches, to our knowledge, developed for an automated brand personality detection task. In this paper, we introduce a fully automated machine learning approach for practitioners and academic scholars to analyze how personalities of brands are channeled and perceived via social media networks, and we offer a foundation for future advances in examining brand-consumer relationships occurring in social media networks.

We intended to integrate closed vocabulary based methods, supervised learning, and unsupervised open vocabulary methods into one refined model. At a high level, our algorithmic design takes the unstructured text data from social media accounts and returns scores for Aaker's (1997) five brand personality dimensions; Sincerity, Excitement, Competence, Ruggedness, and Sophistication in real time. Put concisely, our model provides a novel method of analyzing social media content that may considerably increase the scale and scope of brand research.

PRIOR LITERATURE

In the following paragraphs, we discuss relevant work from both human and brand personality literature while illustrating the computational approaches in each domain.

Computational Methods in Human Personality

Briley and Tucker-Drob (2014) define human personality traits as, "individual differences in general patterns of thoughts, feelings, and behavior". It is widely accepted that human personality exhibits expression on five factors: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness (Goldberg 1990). The social computing research has recently shown interest in predicting human personality from SMN (Chen et al. 2014; Golbeck et al. 2011) and developing personalized systems (Gou, Zhou, & Yang, 2014). Personality prediction task is mainly achieved through content-based analysis on text documents such as essays (Mairesse et al. 2007; Pennebaker & King, 1999) and Facebook and Twitter posts (e.g. Golbeck et al. 2011). Among these content-based studies, some researchers combined social media language usage (Park et al. 2015) with the examinations of distinct features such as age and gender (Schwartz et al. 2013) to conduct automatic human personality assessments. In general, language content and social

network activity have become the most important predictors for human personality assessment in SMN (Markovikj, Gievska, Kosinski, and Stillwell, 2013).

Most of the personality studies, with few exceptions, have utilized a closed-vocabulary word counting approach to analyzing language. This method uses previously formed lists of words and then counts the relative frequency of these words within a body of text. Linguistic Inquiry and Word Count (LIWC: Pennebaker et al. 2007) is a popular implementation of this method. More recently, personality research have developed and implemented finer grained, open-vocabulary methods for language analysis (e.g. Park et al. 2015; Schwartz et al. 2013). Open-vocabulary methods do not rely on priori word judgements; instead, they incorporate unsupervised machine learning models that cluster semantically related words (Blei, Ng, and Jordan, 2003). In addition, open-vocabulary methods can use uncategorized words, nonword symbols (e.g. emoticons), and multiword phrases along with topic clusters to extract a comprehensive collection of language features from the body of text (Park et al. 2015). Several studies compared open and closed vocabulary methods in the context of personality prediction from text. Models using open-vocabulary and machine learning significantly outperformed closed-vocabulary models when predicting the personality of social media users (Iacobelli, Gill, Nowson, & Oberlander, 2011; Schwartz et al. 2013). Thus, these previous studies suggest that adopting advanced analytics in the forms of supervised or unsupervised learning methods may result in finer grained analyses in personality prediction studies.

Computational Methods in Brand Personality

In brand literature, conventional empirical methods including self-reported surveys and standard personality questionnaires have been widely used for data collection and hypotheses testing efforts (e.g. Aaker, 1997; Carr, 1996). The emergence of social media networks, such as

Facebook and Twitter, have created novel online platforms for brands to interact with humans. Such platforms have already transformed consumer behavior in terms of the creation of large amounts of user generated content and mass consumption of this content. In addition, this transformation has generated vast data sources for marketing scholars and practitioners to unlock new consumer insights by using modern data analytic techniques (e.g. Zhang, Bhattacharyya, and Ram, 2016). As a result, the emergence of social media networks not only provides unbounded data sources to empirically test propositions for various disciplines, but it also enables the implementation of advanced analytical methods that considerably enhance the scope and scale of personality research (Golbeck, Robles, Edmondson, & Turner, 2011). We note that there is a need for such analytical advancements to be applied to the realm of brand personality, to assess how brands personalities are being channeled and how they are being perceived in people's minds (Aaker, 2012). The work by Xu et al. (2016), possibly the most related work in this context, conducted a predictive analysis on the drivers of brand personality embodied in social media. The authors focused on the factors that drive brand personality instead of direct brand personality prediction from the social media content. They used questionnaires and a closed-vocabulary approach (LIWC) as an illustration of the consumer-perceived brand personality without employing machine learning and advanced analytic implementations such as open-vocabulary based approaches (e.g. unsupervised cluster detection) and social network analytics (e.g. link prediction in a social network).

Although it is relatively rare, we have observed a growing interest recently in social media analytics implementations of machine learning within the realm of marketing research. For example, Culotta and Cutlar (2016) developed an automated data analytics tool to predict brand perceptions from Twitter. Similarly, Huang & Loa (2016) applied supervised machine learning to

elicit consumer preferences. In addition, Jacobs et al. (2016) integrated unsupervised learning for better identification of items purchased together.

However, in the context of brand research, to our knowledge, there are no extant machine learning approaches developed for brand personality prediction. To provide a clear demonstration about the positioning of our work amongst current literature, we categorize personality research along two dimensions: type of methods and domains of analysis. We classified the type of methods as conventional and social media analytic methods. The conventional methods column refers to methods that do not use machine learning and social media analytics in personality detection, such as self-reported surveys, questionnaires, and closed-vocabulary based static linguistic methods. The social media analytics column refers to automated methods utilizing machine learning and other advanced analytic implementations such as open-vocabulary based approaches (e.g. unsupervised cluster detection) and social network analytics (e.g. link prediction in a social network). Figure 3.1 depicts our contribution within the realm of human personality and brand personality research.

	Conventional Methods	Social Media Analytics
Human Personality	☑	☑
Brand Personality	☑	**Our Research (Machine Learning)**

Figure 3.1: Personality research across methodological approaches

THEORETICAL BACKGROUND: BRAND PERSONALITY

The term brand personality, first coined by Martineau (1958) who proposed that consumer behavior is dependent upon personality rather than objective reality by referring to a set of human characteristics related to a brand. For instance, users have characterized the brand personality of

Mercedes Benz as *upscale* and *aspirational,* while Calvin Klein's brand personality has been characterized as *sexy* and *sophisticated*. There are product-related and non-product-related factors that drive the formation and perception of brand personality (Aaker, 2012). On one hand, the attributes of a product can signal a personality such that a high-priced Burberry scarf might portray signals of wealth, style, and perhaps a bit arrogance. On the other hand, non-product-related signals can include age, symbols, employees, CEOs, celebrity endorsers, and sponsorships. For instance, the targeted sponsorship of the International Ice Skating Championship may reinforce the Red Bull's *offbeat* and *youthful* personality. Considering all of these factors with brand personality formation, the growth of SMN have sparked opportunity in how firms foster their branding efforts by performing integrated marketing activities with much less effort and cost than before. More specifically, firms can form and channel a composite brand personality through SMN by utilizing efficient and multifaceted communication in everyday conversations.

Once properly formed, brand personality can be an eminent asset for firms. Personification of brands may provide an important point of differentiation from competitors, and assist corporations in developing brand equity (Ross, 2008). Since the concept of brand personality emerged over decades ago, there has been a growing interest in the subject among scholars (Aaker, 1997; Carr, 1996). Marketers therefore need to ensure that a brand's personality is channeled consistently to the consumers. When a brand consistently nurtures its brand personality, the relationships between the brand and its consumers evolves in a way that is characterized by the values inherent in the brand's personality (Fournier, 1998).

Corporate brands exhibit brand personality that represents various characteristics of the brand, and this personality evolves largely from the brand's fundamental values and positioning (Harris & de Chernatony, 2001). The goal of corporate branding efforts is to develop a brand which

is perceived as unique and of high value (Keller & Lehmann, 2006). Consumers' perceptions and behaviors are influenced by the brand personality that is channeled from the focal firm.

To date, Aaker's (1997) brand personality scale is the most widely employed brand personality measure for theoretical understanding of the brand personality construct. She analyzed the individual ratings of 37 brands on 114 personality traits by 613 respondents from the United States and developed a reliable, valid and generalizable scale to measure brand personality (Aaker 1997). As a result, brand personality scales are composed of 42 traits. These traits are defined into five dimensions: Sincerity, Excitement, Competence, Sophistication and Ruggedness. Sincerity captures traits such as down-to-earth, cheerful, sincere, and friendly. Excitement indicates traits including daring, young, trendy, imaginative, unique and independent. Competence is represented by traits such as intelligent, reliable, secure, and successful. Sophistication is characterized by traits including upper-class, glamorous, charming, and good-looking. Finally, ruggedness encapsulates traits such as masculine, tough, and outdoorsy.

SMNs add another layer on top of branding strategies, as social media offers brands an option to have a literal voice that speaks directly to consumers in everyday conversations. Interaction between brands and consumers through social media is persistent and visible to all, and social media serves as a mirror to reflect "public displays of connections" (Boyd & Ellison, 2007). In addition, SMN offer exponential spread of content, along with its unprecedented accessibility. These features of SMN make it the fastest-growing marketing channel in the world (Evans, 2012).

SMN content enables the analysis of both channeled and perceived personality of a brand. On one hand, previous theoretical work on brand personality formation suggests that consumer-perceived and employee-perceived brand personality have more predictive power than channeled personality (e.g. official social media account announcements) in brand personality formation (Xu

et al. 2016). On the other hand, properly and consistently channeled brand personality has a significant effect on audience perception (Parker, 2009). Congruence between channeled and perceived brand personality have received past research attention and analyzed with congruity theory (Parker, 2009). Table 3.1 provides examples of channeled personality dimensions from official brand SMN account (e.g. Twitter) and perceived personality from user accounts.

Table 3.1. Channeled and perceived brand personality examples based on Aaker (1997)

Brand Personality Dimensions	Channeled Personality	Perceived Personality
Sincerity	"Please join us supporting Pets for Vets, a program dedicated to supporting veterans and providing a second chance for shelter pets by pairing them with America's veterans who are looking for a forever friend. To donate, please visit: #considerate"	"Pleasant surprise! thanks @VirginAmerica for my sweet treat of the day!"
Competence	"Law firms must find ways to operate more efficiently in order to stay competitive."	"Congrats to @McKinsey for ranking among top consultants of choice for achieving intelligent operations @HfSresearch"
Excitement	"How exciting! Finish up the semester with an energizing Red Bull. Have you tried one before?"	"@redbull If im ever sponsored by you guys i will do something crazy."
Ruggedness	"The outdoor routes don't stop if you don't. #justdoit"	"this choker trend is wild y'all i wore a shoelace to the bars last night & i've never gotten so many compliments thanks @Nike just do it lol"
Sophistication	"Elegant and seductive lips with sheer understated eyes – the @Burberry siren red runway look:"	"@Burberry snapchat is so stylish and classy"

This work may also be complimentary to several other theoretical studies such as analyzing the impact of specific group perceptions on brand personality. For example, researchers may investigate the relative effect of employees' perceptions on their brand personality by using the employee-generated content on public forums (e.g. Glassdoor.com). Thus, we believe that the introduction of a novel, automated, and highly generalizable method to extract near real-time

estimates of brand personalities from the generated text content may have strong implications for the academic community in enhancing the theoretical understanding of the brand personality construct.

RESEARCH SETTING AND METHODOLOGY

Our methodology was guided by design science research principles to report relevance and enhance the rigor of our research process and results (Peffers et al. 2008). According to Peffers et al. (2008), we first introduce the design and development process in the following sub-sections: Sample selection and data sources, machine learning implementation, clustering, and classification phases. Second, we demonstrate the results. Then we evaluate the results and test robustness. Finally, we conclude the paper with communication and implications.

Sample selection and data sources

To test the generalizability of our approach across brands, we use a wide range of brands from a variety of sectors. To collect brands, we used the website millwardbrown.com, which maintains a large selection of brands categorized by sector including apparel, cars, luxury, personal care, food drink, financial institutions, technology, telecommunication, insurance, and airlines. We trained the algorithm by using five well-known brands from each category that totals up to 50 brand accounts for our learning model. We then used additional 20 brands for testing the results of our framework. These 20 brands for testing were randomly chosen based on their publicly perceived visible personalities from different industries. For example, Virgin America is a brand that signals entertaining, sincere, and authentic personality, whereas financial services firms such as McKinsey signal strong and competent personality. Thus, we assumed demonstration of our results would be

more interpretable if we focus on strong and publicly visible brands in the first stage of testing our algorithm.

For each of these brands, we utilized the Crimson Hexagon data base to collect Twitter and Facebook posts from official brand accounts starting on June 1st, 2014 until June 1st 2016. Crimson Hexagon warehouses all public Twitter and Facebook data stretching back to 2009. Consequently, we retrieved 26,834 posts in total for the training set and 6,388 posts for the test set.

Machine Learning Implementation

Algorithmic classification procedures can be examined under two high level methods, namely supervised and unsupervised learning. In general, supervised models define and explain phenomena which are embedded in the dataset and can be used for predicting the value of the target attribute (in this case these attributes are brand personalities) knowing the values of the input attributes. For supervised learning algorithms, a given data set is typically separated into two parts: training and testing data sets with known category labels. The learning algorithm is applied to the training data and then the training goes through several iterations. After each iteration, the result is updated using labeled input. In contrast, unsupervised models do not require prior training in order to mine the data. For instance, a typical unsupervised text mining algorithm determines which terms or phrases are related and groups them into clusters which especially can be helpful discovering hidden topics embedded in complex data and providing an organized view of the data to facilitate decision making processes. Since supervised methods use training sets as references and blends the knowledge of expert manual coders and algorithmic automation, the output often tends to be more accurate than unsupervised models (Berry et al., 2015).

We adopt a hybrid method that utilizes the strengths of unsupervised learning and incorporates extracted output into supervised learning to examine our theoretical prediction for two main reasons. First, we need the highest possible accuracy on identifying brand personalities from the collected social media corpus which should be close enough to error-free human manual coder performance. Unsupervised methods alleviate the complexity of the data and provide a simpler view for decision makers, but its performance will be only supportive when it comes to exact identification of brand personalities within the large unstructured text data. Second, since the cost of human coder knowledge for labeling a relatively large subsample in the training phase is expensive and not practical to train very large datasets, we focus on utilizing LDA and Word2Vec unsupervised models along with previously identified brand personality word libraries (Aaker, 1997; Opoku et al., 2006) to implement our supervised methods.

Unsupervised techniques are also sometimes used to generate a valuable subsample of data for human coders. For example, a sampling-by-clustering algorithm proposed by Zhu et al. (2008) shows us a way to form an initial data set for the labeling phase. This sampling-by-clustering algorithm overcomes the problem of selecting representative samples. In summary, the entire unlabeled corpus of tweets or Facebook posts is partitioned into a predefined number of clusters. The sampling-by-clustering algorithm uses cosine-based distance measure and K-means clustering to estimate similarity among posts and assign the subsequent posts closest to the centroid of each cluster (See Zhu et al, 2008; Duda & Hart, 1973 for the details of this clustering algorithm). In this study, we first follow the similar yet deeper approach to identify valuable clusters and word representations of each cluster while incorporating the extracted output with previously proposed word dictionary of brand personalities to label the training set. A simple visualization of flow diagram of this Phase-1 process is presented below (Figure 3.2). Ultimate goal of Phase-1 is to

112

achieve a labeled training set without any human intervention. Once we obtain the labeled data, we can apply supervised machine learning methods to assess and classify each document's brand personality in Phase-2. We refer each account as a separate document in this model. Each document composes of thousands of tweets and Facebook posts collected within the specified period. Hence, the unit of analysis of this model is a brand and its correspondent brand personality scores for each account.
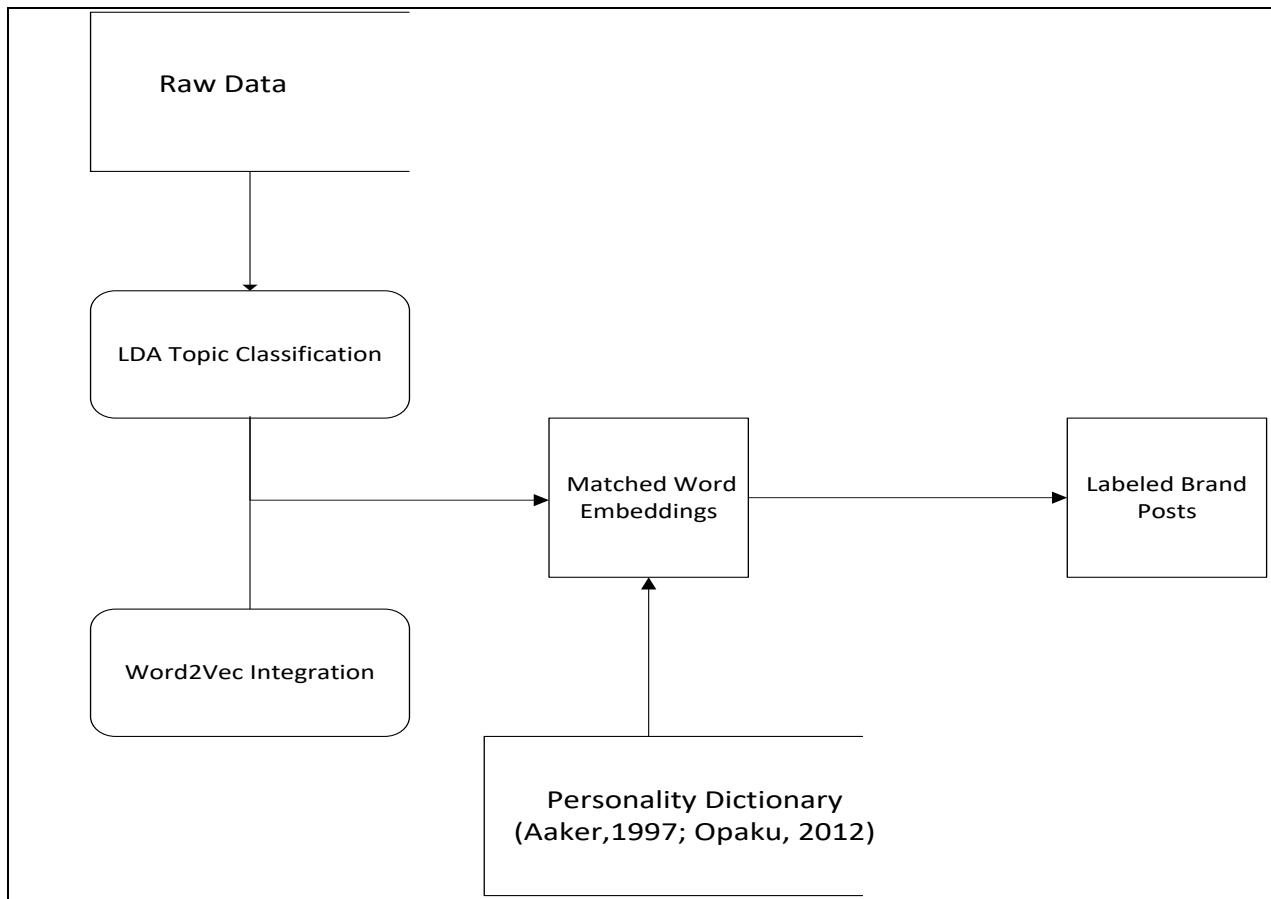

Figure 3.2. Labeling Process (Phase-1)

Phase-1: LDA Topic Clusters and Word2Vec Word Representations

Latent Dirichlet Allocation (LDA) (Blei et al., 2003) is a model that has gained popularity among scholars as a tool for automatic corpus summarization and visualization. LDA is an entirely unsupervised algorithm that models each text document as a mixture of topics. The model

113

generates automatic summaries of topics in terms of a discrete probability distribution over words/terms for each topic, and further understands per-document discrete distributions over topics. LDA makes the explicit assumption that each word is produced from one topic. Although LDA is illustrative enough to generate multiple topics per document, it is not sufficient for multi-labeled corpora because, as an unsupervised bag-of-words model, it offers no obvious way of incorporating a supervised label set into its learning procedure. In brief, LDA models document-word-relationships by using a global bag-of-words approach which disregards local relationships (e.g. word order or grammar) while focusing on the frequencies with which words appear.

To overcome the above-mentioned deficiency of LDA, we incorporate Word2Vec (Mikolov et al., 2013) to leverage both from global and local presentations of terms among clusters. Word2Vec is a predictive algorithm for learning embeddings using a deep neural network model. Embeddings are vector representations of words represented by a set of hidden variables, and each word is represented by a specific embodiment of these variables. Word2Vec directly try to predict a word from its local neighbors in terms of learned small, dense embedding vectors. Concisely, Word2vec tries to overcome the following problem exists in natural language processing implementations.

Traditionally, words are treated as discrete atomic symbols, and therefore 'San Francisco' may be represented as Id001 and 'Los Angeles' as Id999. These encodings are arbitrary, and provide no useful information to the system regarding the relationships that may exist between the individual symbols. This means that the model can leverage very little of what it has learned about 'San Francisco' when it is processing data about 'Los Angeles' (such that they are both cities, they are both in California, and they are both in the West Coast). Representing words as discrete ids subsequently leads to data sparsity. Using word embeddings can overcome some of these obstacles

and refine topic models with relevant and salient terms. For instance, in the following figure, the LDA algorithm provides a list of salient terms in Topic 26 for a courier service document. When we go over the output, although it provides some sense about the topic 'delivery', the rest of the salient terms do not provide an insight to the reader. With incorporating Word2Vec to the model, we would not see the terms 'bigdata', 'voting', 'sylvia' or 'baseball' in this topic (on the y axis of bar chart in Figure 3.3) because these word embeddings infer distinct phenomenon instead of the quality or relevant features about delivery. Thus, by incorporating Word2Vec into the model, we expect to achieve more refined topics contain more relevant words that can provide a cumulated sense to the reader about the topic.
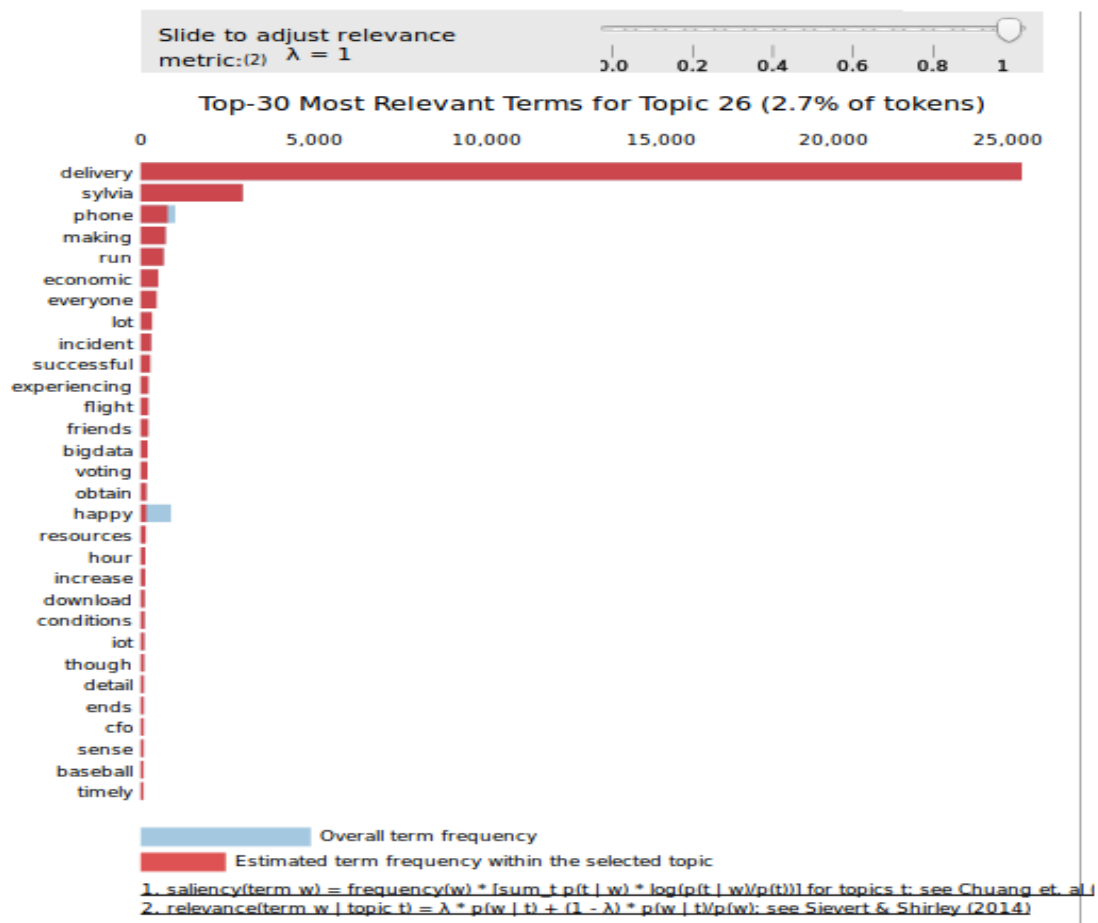

Figure 3.3. Visualization of terms in a sampled topic with LDA

115

Instead of drawing solely on LDA based topic clusters, we integrate Word2Vec into the clustering model and provided output in shown in Figure 3.4. Within the 'delivery' topic, we witness relatively more applicable terms such as 'confidential', 'privacy', 'risk', 'flight', and 'safety' than the output terms presented in Figure 2. Understanding the lexical usage of a word within a document requires not only extracting term global frequency but also deriving local relevance and saliency from word embeddings. The importance of such differentiation is illustrated in this paper's evaluation section, with a real data example that shows the power of this method in increasing the accuracy of the data training/labeling process.
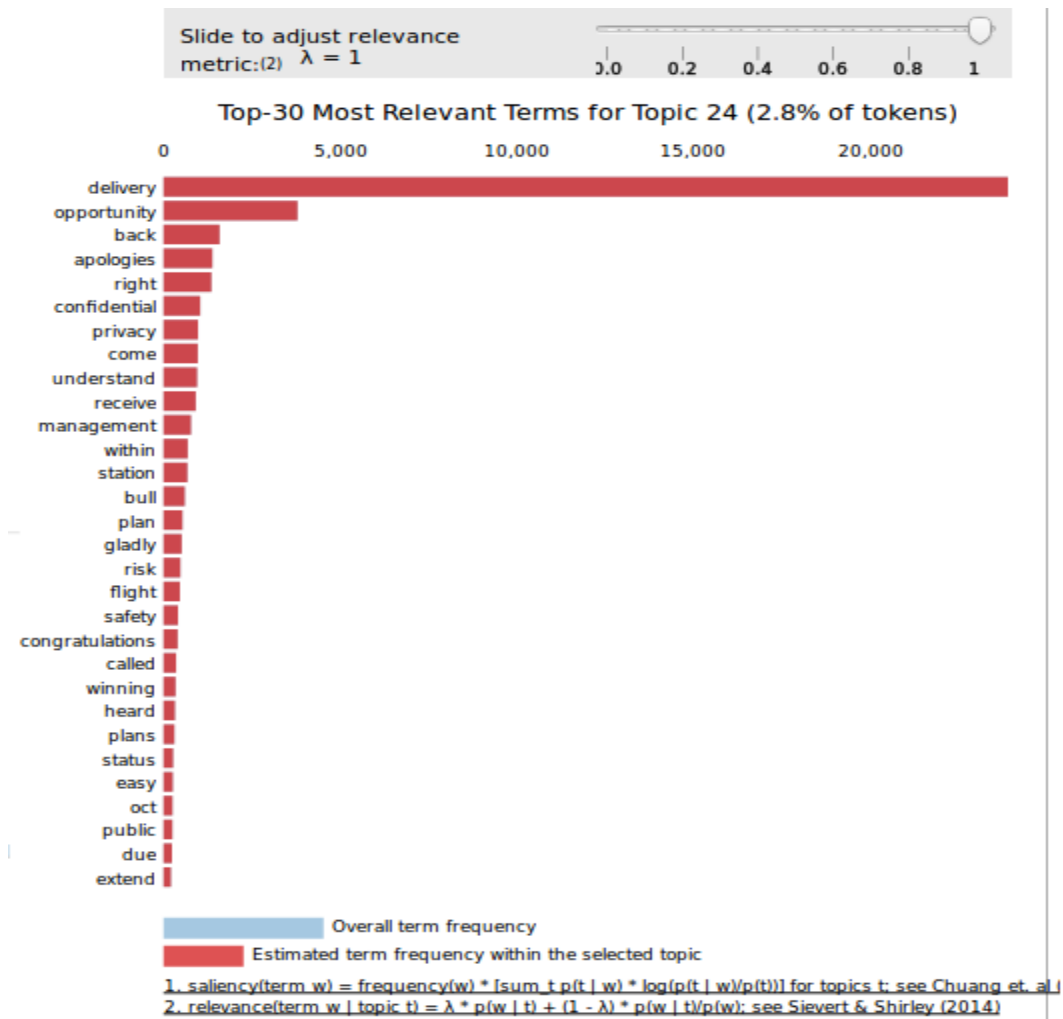

Figure 3.4. Visualization of terms in a sampled topic with LDA & Word2Vec

After implementing Phase-1, we pull relevant words by leveraging Aaker (1997)'s theoretical trait norms from her brand personality dictionary, and we combine these traits norms with the synonyms provided by Opoku's (2006) brand personality dictionary. In other words, we conducted a closed-vocabulary approach for this step. Utilizing this key word list is only partially helpful in the detection of brand personalities due to its lack of comprehensiveness. Thus, we combine the terms from this dictionary with our method of analyzing refined topic clusters to label posts with one of the brand personality classes. Putting it differently, we conducted a version of open-vocabulary approach by extracting topic models and relevant terms in each topic by using hybrid LDA & Word2Vec implementation. Therefore, we were able to label social media posts in specific clusters that matched with word embeddings drawn from previously published brand personality dictionaries. For example, animal and dangerous are two terms in the dictionary infer the *ruggedness* personality dimension. Solely key word-based labeling (also known as closed-vocabulary approach) would fail if it did not see these words in any posts. By leveraging word embeddings, we can label a specific previously unlabeled post with *ruggedness* which includes words such as 'tiger', 'snake', or 'scorpions' since these words also infer dangerous animals, and are identified through our hybrid detection method. Upon completion of this phase, we achieve the labeling of 26,834 posts from different brand accounts into one of the brand personality dimensions – sincerity, ruggedness, competence, excitement, and sophistication.

Phase-2: Classification

After forming our training set, first, we clean up the raw data set by applying pre-processing to remove stop words, stemming, and punctuation, and transform it to a computational format by using scikit-learn machine learning package for the Python programming language (Han, Kamber, & Pei, 2011; Pedregosa et al., 2011). We then conduct feature extraction to transform unstructured

text data into numerical vectors for computational processing. Feature extraction is the process of taking text and splitting it into individual terms. This process takes these sets of terms and transforms them into numerical feature vectors. We leverage existing scikit-learn Python modules to apply feature extraction. Before we move on to the computational details of our learning models, we provide an overview of the steps to detect brand personalities below and illustrate the flow diagram in Figure 3.5:

1. Form the training set by using unsupervised LDA and Word2Vec with brand personality word dictionaries
2. Extract features by using scikit-learn Python modules (e.g. CountVectorizer)
3. Employ supervised machine learning algorithms and validate the model accuracy (e.g. Random Forest Classifier)
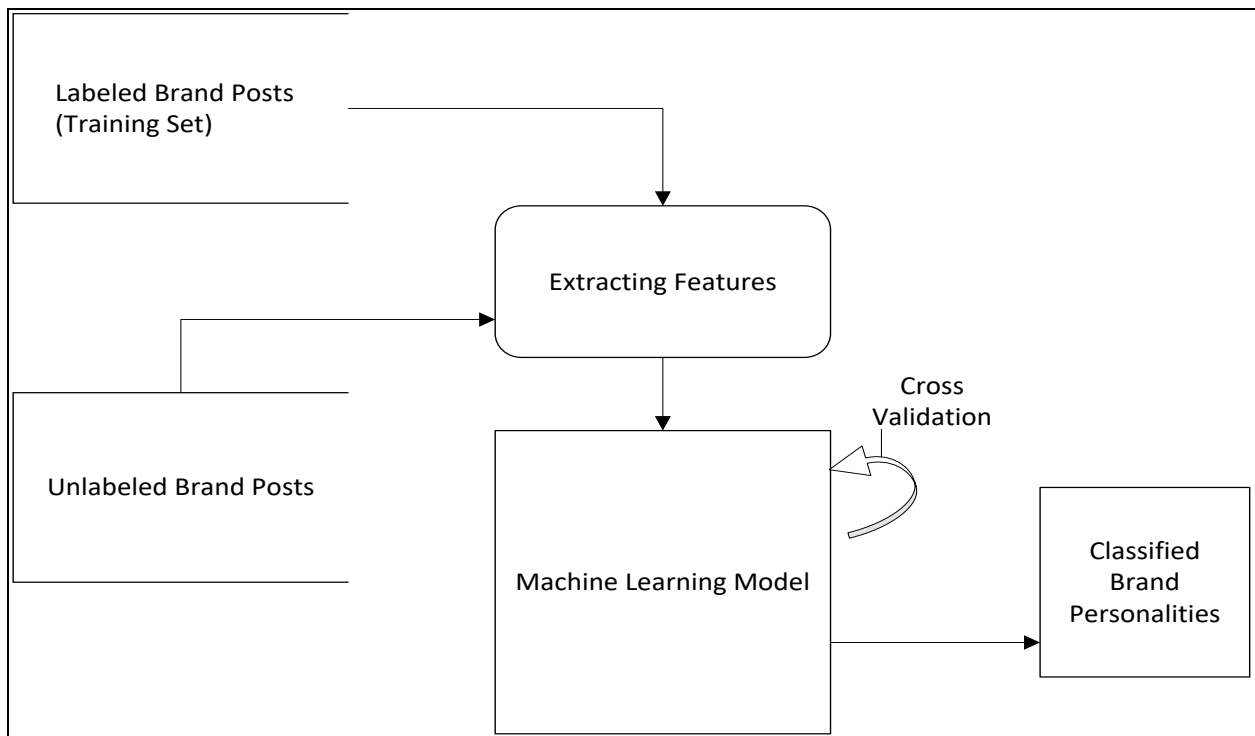4. Test and quantify the personality scores of brands (e.g. Cross Validation)

Figure 3.5. Classification Process (Phase-2)

The third step is the application of machine learning models to predict the personality of a given independent brand. Text classification through supervised learning techniques has increasingly been employed in mainstream information systems literature. The goal of this step is to select the best classification method for our analysis from the alternatives, keeping in mind that our main priority is to minimize classification error and that our context is one where there are multiple classes for prediction (5 brand personality classes).

We examined four classification approaches and used the best performing algorithm in each type of method. The classification methods could be broadly categorized as frequency-based (e.g. Naïve Bayes), proximity-based classifiers (e.g. K-Nearest Neighbor), non-probabilistic linear classifiers (e.g. Support Vector Machines), and decision tree based classifiers (e.g. Random Forest). We observed a 94.34% accuracy rate for the random forest model and 92.38% for the support vector machines (SVM) model. The accuracy levels of each classification method is shown in Table 3.2. Our method for testing the accuracy of each classifier was by using 10-fold cross validation with the pre-labeled training set (26,824).

Table 3.2. Comparison of Classification Algorithms

Machine Learning Classification Method	Accuracy with 10-Fold Cross validation
Random Forest	0.943
Support Vector Machines	0.924
K-Nearest Neighbors	0.874
Naïve Bayes	0.863

DEMONSTRATION

Since we achieved relatively sufficient accuracy (94.3%) with the random forest model and we made sure that there was issue of over-fitting problem by conducting cross validation, we apply this learned model to predict the following 20 brand's personalities. Since one may question the robustness of the labeling process and the reliability of our proposed model, we first aim to

demonstrate the results for 20 brands. Overall findings are presented in Table 3.3. The numbers

corresponding to brands represent the density percentage of each dimension. Note that one post

may include several separate traits and may not always signal an exclusive dimension. Thus, we

employ a weighted scale on each dimension and extract the probabilities for each dimension that

brands engage in. We discuss the results of each of five dimension in the following subsections.

Table 3.3. Overall Results for the Test Set (20 Brands)

|	COMPETENCE	EXCITEMENT	RUGGEDNESS	SINCERITY	SOPHISTICATION
Deloitte	0.243	0.387	0.074	0.237	0.056
BCG	0.233	0.387	0.076	0.241	0.059
McKinsey	0.213	0.421	0.078	0.225	0.061
Fedex	0.199	0.338	0.038	0.382	0.041
Coca Cola	0.175	0.341	0.053	0.338	0.09
T-Mobile	0.159	0.346	0.079	0.327	0.087
Nike	0.155	0.33	0.131	0.3	0.081
DrPepper	0.146	0.307	0.066	0.38	0.1
Virgin America	0.144	0.324	0.05	0.407	0.073
Jeep	0.141	0.368	0.096	0.297	0.096
Victoria Secret	0.132	0.336	0.057	0.344	0.129
Cabelas Outdoor	0.127	0.318	0.126	0.335	0.091
Red Bull	0.12	0.671	0.089	0.339	0.103
Pepsico	0.117	0.357	0.068	0.329	0.126
Spirit Airlines	0.112	0.351	0.072	0.37	0.092
McDonalds	0.111	0.36	0.068	0.384	0.075
Toms	0.11	0.384	0.061	0.331	0.106
Dove	0.095	0.341	0.032	0.157	0.042
Louis Vuitton	0.094	0.442	0.071	0.209	0.181
Burberry	0.058	0.448	0.055	0.204	0.233

The numbers refer to percentage score of each dimension per brand

1. Sincerity

To provide a detailed analysis of our extracted results, we analyze each dimension one by

one and illustrate the prominent brands who display that specific brand personality. For example,

we provide the sincerity dimension results in Table 3.4. At a high level, sincerity is associated

with traits including words such as domestic, honest, genuine, and cheerful (Aaker, 1997). As

we can observe, Virgin America reflects the highest sincerity among 20 brands, and McDonalds and Fedex come closely after Virgin America.

Table 3.4. Top Scored Brands with Sincerity Dimension
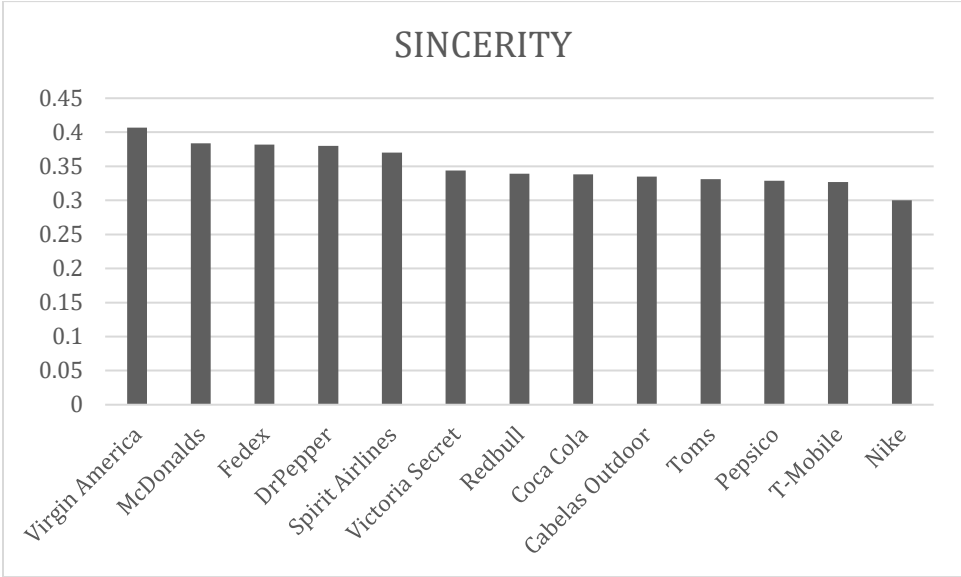
Mean: 0.35, Std: 0.029

The results for sincerity are consistent with channeled personalities of brands. When we look at the corporate webpage and commercials of Virgin America, a strong highlighted personality dimension is sincerity ("make flying fun again"). Virgin America is successfully channeling this personality dimension through social media posts, as our model detected this to be the case. Fullsurge, a strategic consulting firm, published a report(3) about brand strategy and addressed the following statement about Virgin America Consistent with dominant market perceptions of Virgin America, our results indicate that Virgin America successfully channels its core brand personality –Sincerity; cheerful, authentic and down-to-earth—through efficient social media engagement.

3 http://www.fullsurge.com/blog/virgin-america-lesson-brand-personality

"Take Virgin America, for example. The very entertaining safety video, the charismatic and even hip flight attendants, the on-board mood lighting and the way in which you order in-flight movies and menu items set the airline apart. If you have ever taken a Virgin America flight, you know that these touchpoints and interactions are distinctly Virgin. The personality that comes through is strong, unique and authentic."

In addition, we observe strong sincerity for McDonalds, which is actually consistent with the statements about the recently appointed CEO, Steve Easterbrook's, down-to-earth signals and transparent marketing strategy:

"As Easterbrook readies to take the McDonald's helm on March 1 the company has already adopted some of his approaches more widely. Its 'Our Food, Your Questions' U.S. site has 20 million hits on YouTube, addressing queries of customers"

2. Ruggedness

Table 3.5. Top Scored Brands with Ruggedness Dimension
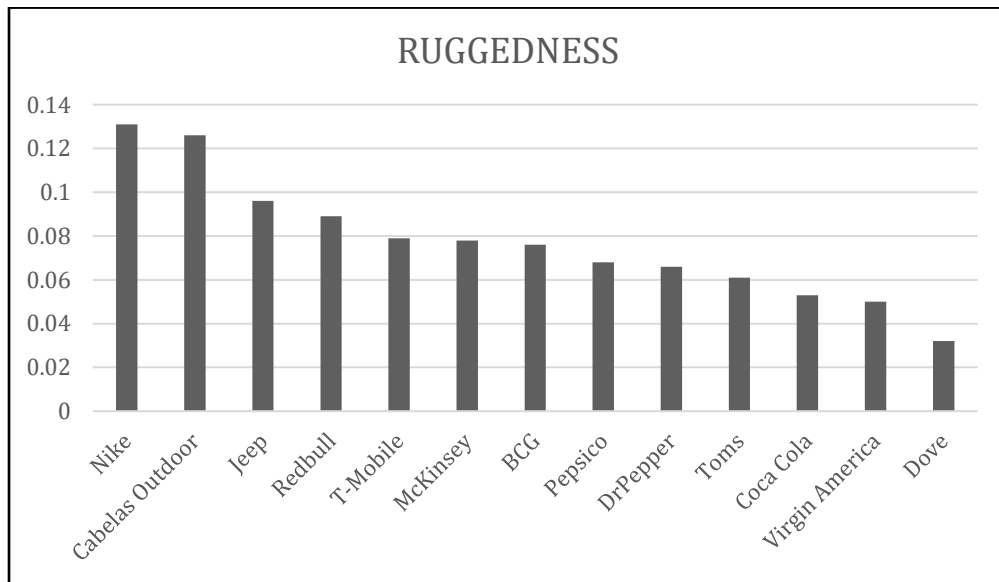

Mean: 0.072, Std: 0.024

The ruggedness personality seems to signal trait words such as challenge, endeavor or outdoorsy trait norms; Nike, Cabelas and Jeep brands are leading in this dimension. Compared to other brands in the sample, these brands are dominantly channeling their core brand personality

through social media channels (Twitter and Facebook). For example, Jeep is considered as a pioneered brand (Bailey, 2016) and clearly signals ruggedness as in the following statement:

"Pioneer brands champion values such as freedom, adventure, self-discovery, self-reliance and ambition. Good examples of pioneer brands are The Discovery Channel and Jeep."

3. Sophistication

Table 3.6. Top Scored Brands with Sophistication Dimension

Mean: 0.096, Std: 0.044
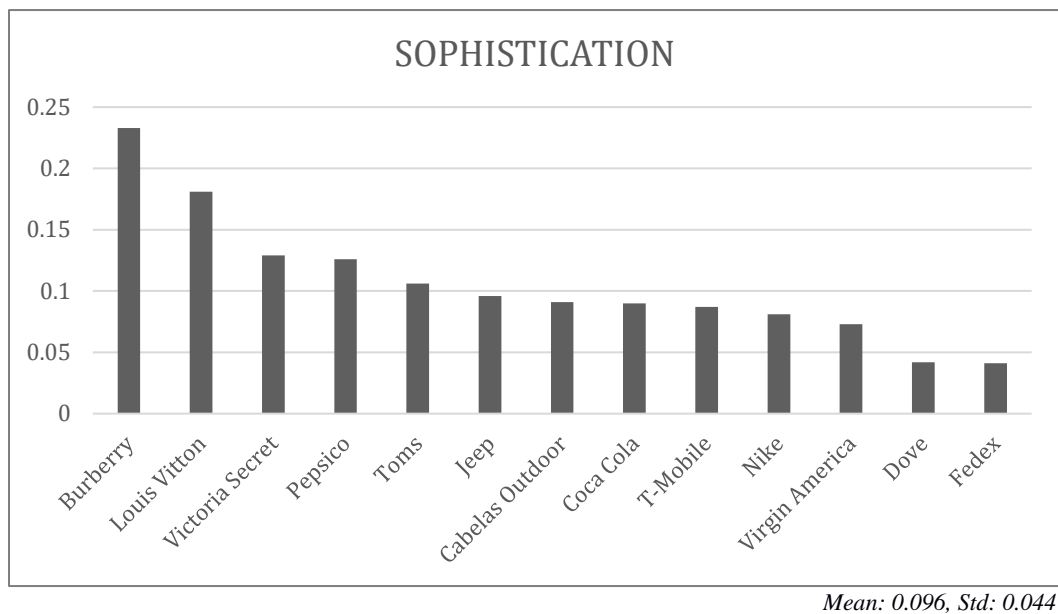
As can be seen above chart, the top three brands that are most strongly exhibiting the sophistication personality dimension – categorized by words such as charming and glamorous (Aaker, 1997) - are Burberry, Louis Vuitton and Victoria's Secret. These results are not that surprising considering how highly these brands produce luxury signals in comparison with others.

4. Competence

Table 3.7. Top Scored Brands with Competence Dimension
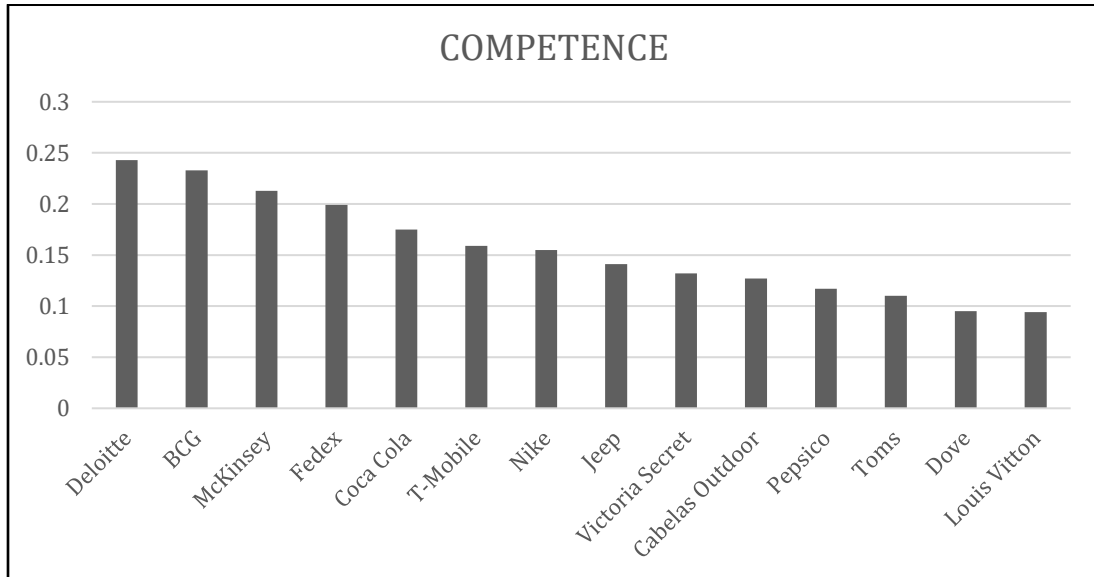
<div align="right">Mean: 0.146, Std: 0.019</div>

Hard working, leader, intelligent, responsible, and proud are some of the trait words that describe the competence personality dimension. The top 3 performers in our analysis are the big consultant companies, which is actually pretty aligned with their mission and vision statements. In addition to their real-life perceptions, we observe a successful channeling of competence personality through social media from these highly competent brands. In addition, Fedex successfully signals its core personality as "reliable and dependable service," and this is even addressed in their mission statement "Safety will be the first consideration in all operations. Corporate activities will be conducted to the highest ethical and professional standards." This is consistent with competence dimension trait norms; reliable, responsible, and dependable (Aaker, 1997).

5. Excitement

Table 3.8. Top Scored Brands with Excitement Dimension

Mean: 0.377, Std: 0.077
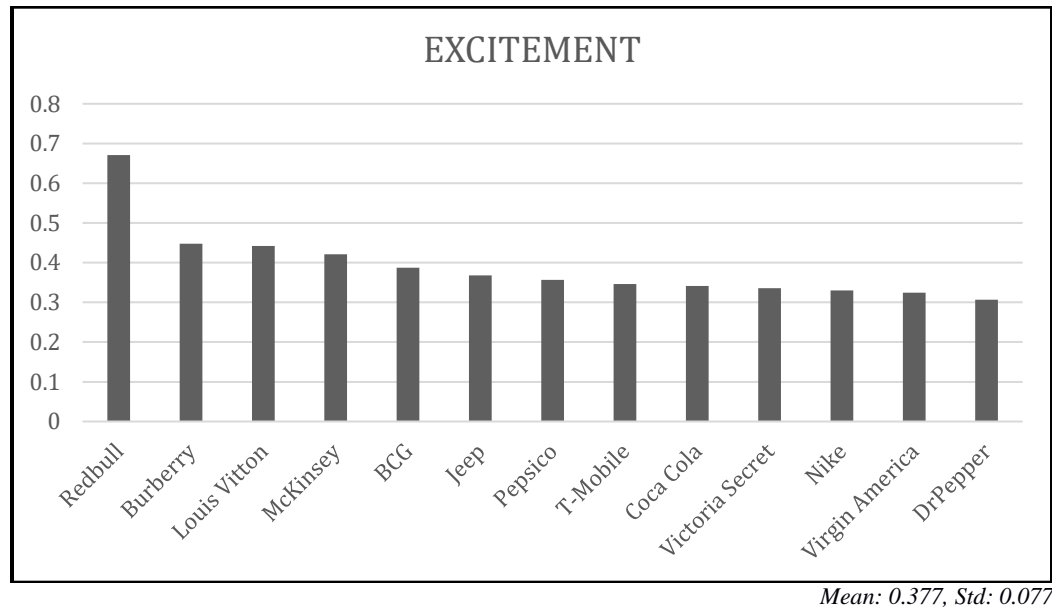
Red Bull is leading in this dimension. Red Bull is an energy drink brand and has a widely known slogan, "it gives you wings" (which brought about a $13 million class action lawsuit in late 2014). This brand slogan presents a signal of excitement and shows our results to be consistent with the general public perception about the personality of this brand. We also found a focus from Burberry on this dimension along with their sophisticated signals.

Additional Demonstration

We also tested our model on firms within the same industry. We selected State Farm, Allstate, and Geico as one of the most prominent brands in the US in insurance industry. Our prior knowledge about the personality of these brands provided as follows:

- State Farm: "Like a good neighbor, State Farm is there" (down-to-earth, friendly, genuine)
- Allstate: "You are in good hands with Allstate" (caring, genuine, tender)
- Geico: "Save 15% or more on car insurance!" (aggressive, audacious, bold, brave)

As we can see from Table 3.9, all three companies successfully channel their core personality as stated above. On one hand, State Farm and Allstate are signaling traits/words such as being down-to-earth and genuine, which represents the sincerity dimension (Aaker, 1997). On the other hand, Geico is signaling daring and aggressive trait norms, which represents the excitement dimension (Aaker, 1997).

Table 3.9. Brand Personality Scores of Insurance Companies

State Farm				
Competence	Excitement	Ruggedness	**Sincerity**	Sophistication
0.09	0.29	0.06	**0.49**	0.07
Allstate				
Competence	Excitement	Ruggedness	**Sincerity**	Sophistication
0.05	0.32	0.04	**0.58**	0.01
Geico				
Competence	**Excitement**	Ruggedness	Sincerity	Sophistication
0.1	**0.52**	0.04	0.32	0.01

Perceived vs. Channeled Brand Personality

As we discussed in the theoretical background section and presented some examples in Table 3.1, our approach enables the analysis of both channeled and perceived personality of a brand. Thus, we additionally collected the conversations of five brands' from Facebook and Twitter. We selected Nike, Virgin America, Deloitte, Red Bull, and Burberry since they channeled the strongest personality for each dimension in our sample set. Then, we ran our algorithm on posts from both

brands' accounts and brands' follower accounts who mentioned that brand and derived the personality scores for each specific brand. Finally, as a demonstration for future theoretical studies, we compared the congruence between channeled and perceived personality by using widely employed similarity metrics, namely cosine similarity.

Cosine Similarity

The cosine similarity measure is widely used to capture the similarity between two vectors in various fields ranging from marketing literature (e.g. Hwang, Bronnenberg, & Thomadsen, 2010), finance literature (e.g. Sabau, 2012), and to information systems literature (e.g., Breese et al. 1998; Salton & McGill, 1986). Cosine similarity takes two vectors and calculates the cosine of the angle between the two vectors as in the following equation where θ is the angle between vectors, the numerator is the dot product of the two vectors, and the denominator is the product of the vector lengths.

Since these vectors only contain non-negative values, the value of the cosine similarity will be between 0 and 1, where 0 represents two completely orthogonal vectors (completely dissimilar) or 1, which represents the same vector (completely the same). In this study, our vectors are personality dimensions that form a vector such as [99, 99, 62, 12, 12]. This vector corresponds to 99% Competence, 99% Sincerity, 62% Ruggedness, 12% Sophistication, and 12% Excitement. To calculate cosine similarity, we utilized the scikit-learn Python package's cosine_similarity function (Pedregosa et al. 2011) and presented three brands' similarity between perceived and channeled brand personality.

Table 3.10. Congruence between perceived and channeled brand personality

Brand	Similarity Distance
Nike	0.915
Virgin America	0.822
Burberry	0.796
Deloitte	0.671
Red Bull	0.611

We show the congruence scores in Table 3.10. Instead of discussing these results, we rather aim to provide a road map for theoretical studies who may examine a research question such as analyzing whether greater congruence between a brand's channeled personality and perceived personality builds a greater emotional brand attachment. If so, what might be the other implications for the strength of congruence between both sides?

EVALUATION AND ROBUSTNESS TEST

To test the generalizability of our approach across different personality dimensions, we considered three alternative labeling techniques and monitored the prediction results. Although our sample results mostly match actual perceptions as illustrated above section, we needed further validation for the accuracy of labeling process. The next step in our process was to compare the labeling inputs which subsequently fed into the subsequent predictive machine learning models. To do this, we compared the LDA & Word2Vec results with Aaker's (1997) trait norms dictionary and human coder evaluations for the sample posts of each brand as described below.

We randomly picked 50 posts for each personality dimension based on LDA & Word2Vec results, which totaled 250 posts to be validated. In addition, we selected 50 more posts which seemed likely to be classified with one of the personality dimensions. Note that, these additional 50 posts initially were not labelled by the automated process. As a result, we extracted 300 posts to compare against methods at the end of first step.

The second step was development of a training document for human coders and a coding scheme to classify tweets into personality dimensions. Morris (1994) tested the validity and reliability of manual coding approaches and achieved an acceptable level of semantic validity. We follow her structural procedure to classify the content based on a coding scheme and to make the results replicable by others. We define single posts as the unit for analysis because they can be objectively recognized by the coders without losing contextual information (Harwood & Gary, 2003).

Next, two research assistants from the authors' institution were initially trained based on the theoretical foundations (Aaker, 1997) and the comprehensive trait norms dictionary (Opoku et al., 2006). The research assistants coded the posts for each of the five personality dimensions.

Several iterative practice sessions were conducted with Twitter and Facebook data sub-samples to train the coders with the content. This sub-sample posts was only used for training of human coders and were eventually excluded from the final dataset. We observed an inter-coder reliability score of 0.89 which is greater than the threshold recommended by Krippendorff (2012). We completed the initial brand personality identification phase by manually training 300 messages with corresponding dimensions. A snapshot of the spearman correlations at this phase is presented in Table 3.11.

Table 3.11. Spearman Correlations of three labelling methods (n=300)

	LDA & Word2Vec	Aaker's dictionary (1997)	Human Coder
LDA & Word2Vec	1		
Aaker's dictionary (1997)	0.51	1	
Human Coder	0.84	0.46	1

As we expected, when we use human coders' ratings as a benchmark, Aaker's key word-based dictionary (closed-vocabulary approach) was not comprehensive enough to assess the personality from a given text. We also observed some mismatch between our LDA & Word2Vec method and human coders' ratings as well. For example, our approach failed to label "@VirginAmerica is the Michael Jordan of airlines" tweet while human coders labelled it as a competence trait. Word2Vec module due not mainly include specific name distances such as Michael Jordan and our approach, as expected, misplaced this tweet within clusters. However, for labeling task, our approach scored higher than a closed-vocabulary based approach and performed almost as good as human coders when there is not specific name dominancy in the text context.

We compiled some examples for both labelling and prediction results in Table 3.12. We provide 10 ruggedness related posts labelled by three alternative methods and the prediction results

of machine learning algorithm (random forest). We chose the ruggedness dimension in our coding sample because the highest intercoder reliability was found for this dimension.

In the first row, the keyword 'animal' is included in Aaker's (1997) trait norms dictionary, which signals ruggedness. Thus, it is labeled as 1. Similarly, LDA & Word2Vec classified it as 1 since this method already comprises the trait norms dictionary. Human coders were also coded this post as ruggedness based on the training knowledge. And our prediction algorithm which provided more than 94% accuracy classified it as ruggedness.

The second post failed to be coded as ruggedness since GoPro is not included in Aaker's dictionary. Similarly, our word embeddings approach failed to compute the distance of GoPro word to other ruggedness related norms in clusters mostly because GoPro is a generic word. However, human coders were aware of its usage in outdoor activities and thus coded as ruggedness.

The third post includes the word 'camping'. Although it is not located in Aaker's list, 'camp' root word is close to other related ruggedness traits in terms of lexical distance computed by LDA & Word2Vec clusters. Human coders also related the word 'camping' with ruggedness. The rest of the lines follow the same logic and output labels are presented.

Table 3.12. Ruggedness personality coding sample

Twitter and Facebook Posts of Cabelas	Aaker's (1997) Dictionary	LDA & Word2Vec	Human Coders	Machine Learning
1. RT @alexis_gass I've never seen so many mounted animals. #Cabelas	1	1	1	1
2. Tag photos of your Favorite Story to #ItsInMyNature to be entered to win a $100 gift card and a GoPro.	0	0	1	0
3. Get to Cablea's Let's Go Camping Workshop today! Learn how to set up a tent and make tasty camping snacks: http://bit.ly/1JtFn6P	0	1	1	1
4. @ad_diamond We love those smiles - even the wildlife approves. #photobomb ^HO	1	1	1	1
5. Commemorate the 50th anniversary of Buck Knives with the Cabela's-exclusive 110 Folding Hunter Knife: http://bit.ly/UQGh78	1	1	1	1
6. The @Gerber_Gear Freescape is a fantastic hybrid blend between camp and kitchen knife: http://t.co/Ptygz5fqlw	0	1	1	1
7. @kayleighs_mom08 We're always excited when a kid is eager to learn about the outdoors. We hope she has a great time!	1	1	1	1
8. Text WARRIOR to 247365 to enter to win 2 free tickets to Warrior Dash. #escapetothedash	0	1	1	1
9. Cabela's 2015 Christmas Sale makes you the hero this holiday season with our Christmas Deals. Find hunting, camping, fishing & outdoor gifts.	1	1	1	1
10. Tell her Congrats! That's a nice fish. ^HO	0	1	0-1	1

One interesting finding is worthy to be highlighted. In the 9th and 10th rows, the word 'fish' is detected. In row 9, the same post includes 'hunting' and 'outdoor', thus Aaker's list labeled

it as 1. Similarly, LDA & Word2Vec and human coders interpreted that post as ruggedness. However, the 10th row is a complicated post as even human coders were confused with regards to its proper labeling. One coder interpreted word "fish" in a restaurant setting and indicated no relationship with ruggedness. The other coder related it with ruggedness since the word "congratulations" used in the same context implied outdoor fishing. More interestingly, our proposed LDA & Word2Vec model labelled it as ruggedness (which is true since the picture provided with that post shows a fishing activity) because of the location of that word in one of the clusters and proximity of that word's distance to other outdoor activity related content embeddings. We examined this LDA & Word2Vec process in detail as presented below.
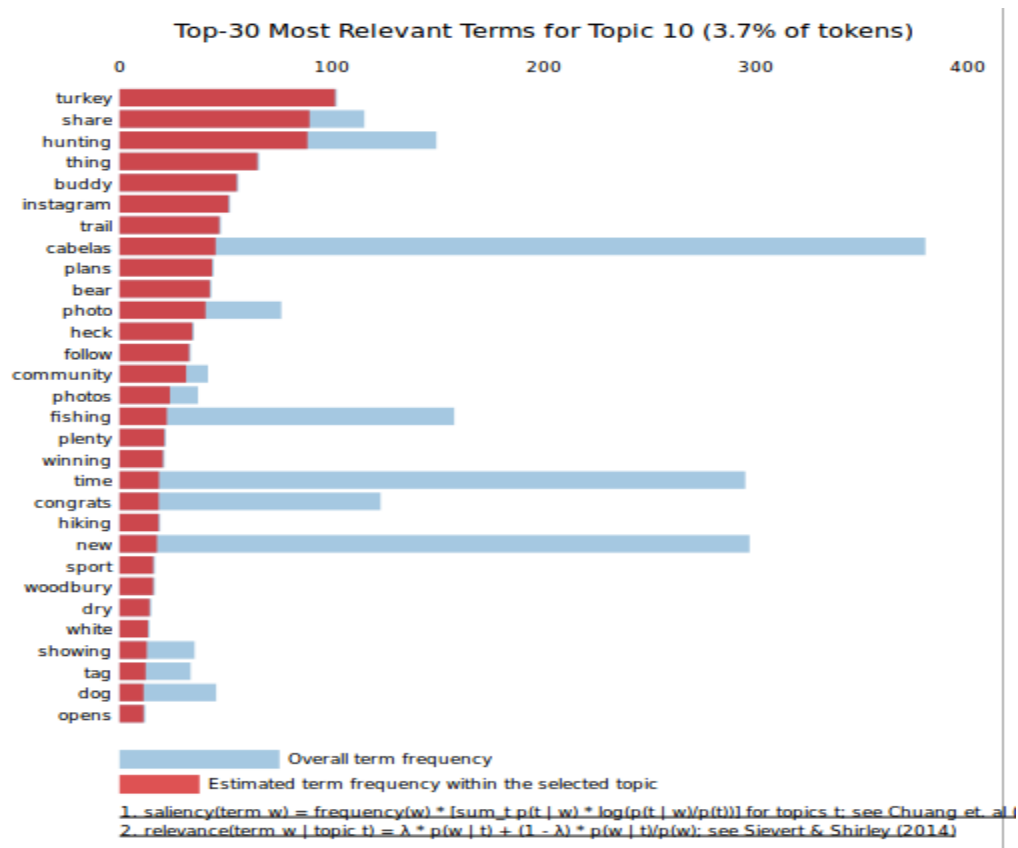

Figure 3.6. Visualization of term 'fish' in a sampled topic with LDA & Word2Vec

Figure 3.6 highlights that animal root words such as 'turkey', 'bear', 'dog', and 'fish' ('fishing' stemmed down to 'fish') clustered into an outdoor activity related context that includes the activities of hunting, sports, and hiking. Thus, although there are insufficient clues for human coders as to whether the word fish was being used in an activity cluster context, our LDA & Word2Vec implementation assigned that post as ruggedness correctly.

Overall, correlation between LDA & Word2Vec and human coder labeling was 84%. Bear in mind that we only compared the agreed posts of human coders with LDA & Word2Vec. In addition, Aaker's (1997) dictionary and human coders showed 46% percent consistency.

Limitations

We point out several limitations with this work. First, it is limited to analyze brands that maintain a Twitter and Facebook presence. Although we have found that all the brands in our sample have a social media presence, there may still be inactivity by some brands which restricts the application of this automated method. Second, data sizes of training sets is the main problem of the machine learning field. While we incorporate a wide range of brands from different industries, clusters may provide better results with more diverse contexts. Third, we rely on text data provided by brands. Future research may also focus on images to identify the real context of the posts instead of solely relying on usage context of word embeddings. However, we hope that the method introduced in this paper provides a useful tool for researchers and practitioners interested in automatically monitoring brand personalities.

COMMUNICATION

A brand personality is an aspect to which a stakeholder can relate. It helps delineate the character of a brand and facilitates the emotional connection between a brand and its target audience. Today's firms are challenged to efficiently define, manage, and control their own brand personality

to succeed a competitive advantage over competitors (Madden, Fehle, & Fournier, 2006). Brand personality can be an influential tool to induce emotions, build trust and loyalty (Fournier, 1998), and enhance consumer preference (Aaker, 1997). To the best of our knowledge, our proposed approach is pioneering. It provides several opportunities for both researchers and practitioners especially in generating personality assessments regularly, easily, and efficiently for brands of interest.

Implications for Researchers

Theoretical studies may expand the five dimensionality of brand personality proposed by Aaker (1997). By drawing on this proposed methodology, new set of dimensions can be detected. There are already discussions about the limitations and generalizability of Aaker's (1997) work (e.g. Austin et al. 2003). By grounding this with marketing theory building and testing research, scholars may test their propositions with our model and provide new constructs and findings to the brand personality literature.

Second, aside from marketing research, proposed model offers several opportunities for information system researchers. Human personality dimensions and brand personality dimensions are treated distinctly in literature. Briley and Tucker-Drob (2014) define human personality traits as, "individual differences in general patterns of thoughts, feelings, and behavior". It is widely accepted that human personality exhibits expression on five factors: Extraversion, Agreeableness, Conscientiousness (or Dependability), Emotional Stability (or Neuroticism), and Culture (or Openness) (Goldberg, 1990). Aaker (1997) simply states that brand personality, "refers to the set of human characteristics associated with a brand". Fournier's (1998) study on consumers and their relationships with brands helped give evidence that there is indeed a relationship between brands

135

and people. Clearly, the line between human personality and brand personality is blurred, and in some senses equal to one another. In the realm of social media, the line between human personality and brand personality is blurred even further, as many brand social media accounts are managed by a single social media manager. So the question is; if a human controls the social media accounts of a brand, how we can better examine brand personality phenomenon considering posted messages generated by a human in a social media environment? We coin an interesting term "IT Personality" which may refer to the behaviors and personalities of humans represented in social media environment. Information systems scholars may theoretically investigate the boundaries between human, brand and IT personality. Since consumers seek brands with personalities that are congruent with either their own or their ideal personalities, do they also exhibit "ideal" or "desired" personality in social media instead of their own personalities? If not, how account managers can successfully reflect the personality of the focal brand that they represent in social media while suppressing their own human personalities? All these questions are worth examining to further understand the effect of underlying IT artifacts on self-presentation in online platforms.

We already discussed the potential implications of channeled and perceived personality congruence in theoretical background section. In the context of congruency, strategy, finance, and organization science researchers may also, for example, examine the congruence between a brand's channeled brand personality and their own CEO brand personality, and analyze the impacts of possible incongruence on firm practices. Furthermore, scholars may look at longtituional data across multiple years and investigate the effects of strategic changes within firms (e.g. CEO turnover) on brand personality over time.

Implications for Practitioners

The use-cases of this proposed method is voluminous. Most directly, marketing managers can monitor how efficiently their brand's personality is being channeled through social media. Since branding strategies can be improved through observed personality and consumer engagement, congruence between channeled and perceived brand personality can also serve as an important metric to evaluate branding strategies. For instance, the impact of dissident stakeholder perceptions on brand personality can be investigated (e.g. employees). Managers even consider reaching out to previously unidentified customer segments by using personality similarity of the brand and social media users. Finally, one specific application of the proposed method example can be identifying celebrities for brand marketing. Since the celebrities are considered to have their own personal brands, measuring the cosine similarity between two may lead practitioners to have greater insight in choosing the most suitable endorser for their brands.

REFERENCES

Aaker, J.L., 1997. "Dimensions of brand personality" *Journal of Marketing Research*, pp.347-356.

Aaker, D.A., 2012. "*Building strong brands*" Simon and Schuster.

Austin, J.R., Siguaw, J.A. and Mattila, A.S., 2003. "A re-examination of the generalizability of the Aaker brand personality measurement framework" *Journal of Strategic Marketing*, (11:2), pp.77-92.

Bailey, L. 2016. "12 Examples of Brand Personality to Inspire You." http://www.careeraddict.com/12-examples-of-brand-personality-to-inspire-you

Berry, M. W., Mohamed, A. H., & Yap, B. W. 2015. "Soft Computing in Data Science," *Proceedings of First International Conference*, SCDS 2015 (545).

Blei, D.M., Ng, A.Y. and Jordan, M.I., 2003. "Latent dirichlet allocation" *Journal of Machine Learning Research*, 3, pp.993-1022.

Briley, D.A. and Tucker-Drob, E.M., 2014. "Genetic and environmental continuity in personality development: A meta-analysis" *Psychological Bulletin*, (140:5), pp.1303.

Boyd, D., 2007. m., & Ellison, NB (2007). "Social network sites: Definition, history, and scholarship" *Journal of Computer-Mediated Communication*, (13:1).

Breese, J. S., Heckerman D., Kadie, C. 1998. "Empirical analysis of predictive algorithms for collaborative filtering" *Proc. 14th Conf. Uncertainty Artificial Intelligence*, pp. 43–52.

Burk, B., 1994. "Position, Personality, not Price, Should Frame Consumer Message" *Brandweek, September19.*

Carr, S.D., 1996. "The cult of brand personality" *Marketing News*, (30:10), pp.4-9.

Chen, J., Hsieh, G., Mahmud, J.U. and Nichols, J., 2014. "Understanding individuals' personal values from social media word use" *In Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pp. 405-414, ACM.

Culotta, A. and Cutler, J., 2016. "Mining brand perceptions from twitter social networks" *Marketing Science*, (35:3), pp.343-362.

Duda Richard O. & Peter E. Hart. 1973. "Pattern Classification and Scene Analysis" *New York: Wiley.*

Evans, D., 2012. "Social media marketing: An hour a day" *John Wiley & Sons.*

Fournier, S. 1998. "Consumers and their brands: Developing relationship theory in consumer Research" *Journal of Consumer Research*, (24:4), pp. 343-353.

Golbeck, J., Robles, C. and Turner, K., 2011. "Predicting personality with social media" *In CHI'11 extended abstracts on human factors in computing systems*, pp. 253-262, ACM.

Goldberg, L.R., 1990. "An alternative description of personality: the big-five factor structure" *Journal of Personality and Social Psychology*, (59:6), pp.1216.

Gou, L., Zhou, M.X. and Yang, H., 2014. "KnowMe and ShareMe: understanding automatically discovered personality traits from social media and user sharing preferences" *In Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pp. 955-964, ACM.

Han, J., Kamber, M., and Pei, J. 2011. "Data Mining: Concepts and Techniques: Concepts and Techniques" *Waltham, MA: Elsevier.*

Harris, F. and de Chernatony, L., 2001. "Corporate branding and corporate brand performance" *European Journal of marketing*, (35:3), pp. 441-456.

Hwang, M., Bronnenberg, B.J. and Thomadsen, R., 2010. "An empirical analysis of assortment similarities across US supermarkets" *Marketing Science*, (29:5), pp.858-879.

Huang, D. and Luo, L., 2016. "Consumer preference elicitation of complex products using fuzzy support vector machine active learning" *Marketing Science*, (35:3), pp.445-464.

Jacobs, B.J., Donkers, B. and Fok, D., 2016. "Model-based purchase predictions for large assortments" *Marketing Science*, (35:3), pp. 389-404.

Iacobelli, F., Gill, A.J., Nowson, S. and Oberlander, J., 2011. "Large scale personality classification of bloggers" *In Affective computing and intelligent interaction*, pp. 568-577, Springer Berlin Heidelberg.

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

Krippendorff, K. 2012. "Content Analysis: An Introduction to its Methodology" *Thousand Oaks*, CA: Sage.

Levy, S. J. 1985. "Dreams, fairy tales, animals, and cars" *Psychology & Marketing*, (2:2), pp. 67-81.

Madden, T.J., Fehle, F. and Fournier, S., 2006. "Brands matter: An empirical demonstration of the creation of shareholder value through branding" *Journal of the Academy of Marketing Science*, (34:2), pp. 224-235.

Markovikj, D., Gievska, S., Kosinski, M. and Stillwell, D., 2013. "Mining Facebook data for predictive personality modeling" *In Proceedings of the 7th international AAAI conference on Weblogs and Social Media (ICWSM 2013)*, Boston, MA, USA.

Mikolov, T., Chen, K., Corrado, G. and Dean, J., 2013. "Efficient estimation of word representations in vector space" *arXiv preprint arXiv:*1301.3781.

Mairesse, F., Walker, M.A., Mehl, M.R. and Moore, R.K., 2007. "Using linguistic cues for the automatic recognition of personality in conversation and text" *Journal of Artificial Intelligence Research*, 30, pp.457-500.

Martineau, P., 1958. *The personality of the retail store*.

Morris, R. 1994. "Computerized Content Analysis in Management Research: A Demonstration of Advantages and Limitations," *Journal of Management* (20:4), pp. 903-931.

Opoku, R., Abratt, R. and Pitt, L., 2006. "Communicating brand personality: are the websites doing the talking for the top South African business schools?" *Journal of Brand Management,* (14:1), pp. 20-39.

Park, G., Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Kosinski, M., Stillwell, D.J., Ungar, L.H. and Seligman, M.E., 2015. "Automatic personality assessment through social media language" *Journal of Personality and Social Psychology*, (108:6), pp.934.

Parker, B.T., 2009. "A comparison of brand personality and brand user-imagery congruence" *Journal of Consumer Marketing,* (26:3), pp.175-184.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., and Duchesnay, É. 2011. "Scikit-learn: Machine Learning in Python," *The Journal of Machine Learning Research* (12), pp. 2825-2830.

Peffers, K., Tuunanen, T., Rothenberger, M. A. and Chatterjee, S. 2008. "A design science research methodology for information systems research" *Journal of Management Information Systems* (24:3), pp. 45-77.

Pennebaker, J.W. and King, L.A., 1999. "Linguistic styles: language use as an individual difference" *Journal of Personality and Social Psychology*, (77:6), p.1296.

Pennebaker, J.W., Booth, R.J. and Francis, M.E., 2007. Linguistic inquiry and word count: LIWC [Computer software]. *Austin, TX: liwc. net.*

Ross, S.D., 2008. "Assessing the use of the brand personality scale in team sport" *International Journal of Sport Management and Marketing*, (3:1-2), pp. 23-38.

Sabau, A.S. 2012. "Survey of clustering based financial fraud detection research" *Informatica Economica*, (16:1), p.110.

Salton G, McGill MJ.1986. Introduction to modern information retrieval. *McGraw-Hill, New York.*

Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Dziurzynski, L., Ramones, S.M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M.E. and Ungar, L.H., 2013. "Personality, gender, and age in the language of social media: The open-vocabulary approach" *PloS one*, (8:9), pp.73791.

Sirgy, M.J., 1982. "Self-concept in consumer behavior: A critical review" *Journal of Consumer Research*, (9:3), pp. 287-300.

Xu, A., Liu, H., Gou, L., Akkiraju, R., Mahmud, J., Sinha, V., Hu, Y. and Qiao, M., 2016. "Predicting Perceived Brand Personality with Social Media" *In ICWSM*, pp. 436-445.

Zhang, K., Bhattacharyya, S. and Ram, S., 2016. "Large-Scale Network Analysis for Online Social Brand Advertising" *MIS Quarterly*, (40:4), pp. 849-868.

Zhu, J., Wang, H., Yao, T., and Tsou, B. K. 2008. "Active Learning with Sampling by Uncertainty and Density for Word Sense Disambiguation and Text Classification," *Proceedings of COLING'08, pp. 1137–1144.*