Sentiment Analysis in Organizational Work:

Toward an Ontology of People Analytics

Roy Gelbard *
Information Systems Program
Graduate School of Business Administration
Bar-Ilan University, Israel
gelbardr@mail.biu.ac.il

Roni Ramon-Gonen
Information Systems Program
Graduate School of Business Administration
Bar-Ilan University, Israel
roni.ramon@gmail.com

Abraham Carmeli
Tel-Aviv University and University of Surrey
avic@post.tau.ac.il

Ran M. Bittmann
Machine Learning Center
SAP Labs, Israel
ran.bittmann@sap.com

Roman Talyansky
Machine Learning Center
SAP Labs, Israel
roma.talyansky@gmail.com

* Corresponding Author

Sentiment Analysis in Organizational Work:
Towards an Ontology of People Analytics

Abstract

The present paper proposes a conceptual ontology to evaluate human factors by modeling their key performance indicators and defining these indicators' explanatory factors, manifestations and diverse corresponding digital footprints. Our methodology incorporates six main human resource constructs: performance, engagement, leadership, workplace dynamics, organizational developmental support, and learning and knowledge creation. Using sentiment analysis, we introduce a potential way to evaluate several components of the proposed human factors ontology. We use the Enron email corpus as a test case, to demonstrate how digital footprints can predict such phenomena. In so doing, we hope to encourage further research applying data mining techniques to allow real time, less costly and more reliable assessments of human factor patterns and trends.

1. Introduction

Organizations and their managements are constantly looking for ways to assess human resource management activities to manage their resources and capabilities more effectively. Organizations often approach consulting firms and research institutions, or run their own processes to assess human resource practices such as employee commitment, engagement and satisfaction. For example, Gallup® conducts an ongoing study of the American workplace, to assess employee engagement and its influence on individual and organizational performance (Gallup, Inc., 2014). Scholars frequently use survey and experimental studies to explore the motivations of individuals, groups and organizations, and the way they act. Despite the useful knowledge these methods yield, they have several shortcomings. First, they rely on self-reports, namely, information provided by the participants themselves about the questions at hand. For example, towards assessing employee commitment to an organization, researchers often ask for employee responses on a set of questions that measure the degree of their commitment to the organization. Although subjective assessments are widely used in fields such as psychology and management and are useful in assessing, for example, psychological attributes and states, caution is called for in interpreting their results, given certain known limitations (e.g., common method bias; see Brutus, Aguinis & Wassmer, 2013; Brutus, Gill & Duniewicz, 2010; Podsakoff, MacKenzie & Podsakoff, 2012). Second, these methods are naturally resource- and time-consuming, despite recently developed platforms (e.g., Amazon Mechanical Turk) that enable quick access of organizations and scholars to potential respondents. Other platforms such as SurveyMonkey® and Qualtrics® offer online surveys designed to collect and analyze data more effectively. However, the basic methodology has not changed. It requires substantial resources, and is challenged by limitations involved in using subjective data. Third, survey data do not provide real-time assessments. Surveys can take variable lengths of time before the data are collected and analyzed. A key issue concerns the inability of organizations or researchers

to approach potential populations on a frequent basis, thus inhibiting the capacity to trace patterns. In fact, many organizations are even reluctant to authorize researchers to conduct theoretical studies that involve surveying their members more than once a year, not to mention problems associated with using the same subjects for different surveys within a short timeframe (e.g., unreliable data due to an emerging automatic response mode). Furthermore, surveys offer a post-factum snapshot of a particular issue. For example, if an organization administers a survey that assesses its employees' job satisfaction on June 30th, the results are unlikely to be available on the same day, and using surveys to capture changes in job satisfaction on a weekly, monthly, or quarterly basis is in fact impossible.

Once acknowledged, these limitations have prompted large ICT companies to search for more effective modes of collecting and analyzing data on human factors (commonly known as people/workforce/HR management/analytics). In Google®, "people analytics" engages in mining human resource and management performance data. "Project Oxygen", for example, identifies characteristics of which great leaders are made by correlating phrases, words, praises and complaints to performance reviews, feedback surveys and nominations for top manager awards (Davenport, Harris & Shapiro, 2010; Sullivan, 2013). Similarly, IBM® and Kenexa® have each developed "HR analytics" to create a smarter workforce by mining real-time information, in ways such as linking HR planning systems with real-time reports.

Following previous research studies (e.g., Gelbard & Carmeli, 2008; Piazza & Strohmeier, 2013), we propose an ontology to evaluate human factors by modeling their key performance indicators (KPI) and defining the explanatory factors, manifestations and corresponding diverse digital footprints of these KPIs. According to Antoniou and Kehagias (2000, p. 623), "An ontology defines the terminology of a domain: it describes the constructs that constitute the domain, and the relationships between those constructs. Every information system uses its own ontology, either implicitly or explicitly. As applications become increasingly complex, we can

observe a trend towards the explicit representation and management of ontologies." We apply a commonly used methodology to reveal human factors by translating prominent key performance indicators into three sets of sub-layers: explanatory factors, manifestations, and raw sources, as illustrated in Figure 1.

-----   Insert Figure 1 about here   -----

We begin by mapping out an ontology that encompasses six main human resource (HR) constructs: performance, engagement, leadership, workplace dynamics, organizational developmental support, and learning and knowledge creation, and proceed to translate them into measurable digital data. By using data mining techniques and sentiment analysis, we have traced patterns and changes in various human factor activities. Next, we propose an initial illustration – rather than a complete solution – of the feasibility of our new approach, by using the Enron email corpus as a case-in-point, to show how digital footprints can predict vitality and engagement. In so doing, we hope to pave the way for a new research direction, and offer a practical ontology that enables linking digital data with HR practices and performance.

The paper is structured as follows: The next section reviews related studies on workforce analytics and human factors. In section 3 we discuss the suggested ontology, and in Section 4 we apply sentiment analysis to organizational emails. In section 5 we present the results, focusing on workforce vitality and satisfaction that emerge from Enron's Email Corpus. Finally, in section 6 we discuss our findings and present our conclusions.

2.   Related work

2.1   *Workforce analytics*

Workforce analytics research aims to offer advanced tools and methodologies for measuring and improving HR processes and undertakings (Gartner, 2017). Such tools may include enhanced recruitment and placement processes, measuring employee performance, learning and development, benefits and perks, employee maintenance, etc. This goal is achieve by analyzing

data from different sources and using advanced methods.

In the past few years, various studies have explored the use of AI, data mining, machine learning and IoT for various HR purposes, such as candidate selection, employee mood and sentiment analysis, churn prediction and more. Different methods have been used to this end: Correlating job requirements with individual resumés (Yi, Allan & Croft, 2007; Bollinger, Hartdtke & Martin, 2012); Analyzing candidate video-clips (such as provided by HireVue) and identifying characteristics or qualities incompatible with the job; Predicting eventual and actual employee attrition by using prediction algorithms and social media data (Punnoose & Ajit, 2016, Robinson, Sinal & Winter, 2014); Identifying employee moods and emotions such as happiness, surprise, anger, disgust, fear and sadness, by analyzing facial expressions captured by the organization's cameras (facial emotion detection) (Subhashini & Niveditha, 2015); Analyzing voice tones being used (Chan & Eric, 2010); Analyzing sentiments through online employee reviews (Moniz & Jong, 2014) and social media platforms (Costa & Veloso, 2015); Inspecting employee productivity by sensors installed on employee badges (Ara et al., 2011). Such sensors enable identifying movement, tone of voice, speech speed, employee cohesion, etc.; Exploring the effect of social media use on employee performance and motivation (Leftheriotis & Giannakos, 2014); Measuring employee knowledge-sharing by analyzing information shared in social media (Zoonen, Verhoeven & Vliegenthart, 2016) or in organizational intranets (Koriat & Gelbard, 2014; Koriat & Gelbard, 2017).

Numerous studies have dealt with other aspects that might be applicable to workplace research. Although those studies have not dealt with organizational research, they indicate that the relevant technology exists, is accessible, and could be applied to organizations. Such studies include: Assessing age and gender based on a picture (Tian & Chen, 2017; Geng, Zhou & Smith-Miles, 2007; Eidinger, Enbar & Hassner, 2014); Sentiment identification by body motions (Stock, Righart & Gelder, 2007) and tone of voice; Activity profiling through body

motion analysis (Preece & Goulermas, 2009; Diego-Mas & Alcaide-Marzal, 2014); Posture identification (Quwaider & Biswas, 2008; Mattmann et al., 2008); Posture analysis to identify emotions (Rosário & Diógenes, 2014; Rosario & Diógenes, 2016) and measure learners' interest level (Mota & Picard, 2003); Analysis of sitting positions by sensors installed on chairs. This enables better adjustment of the chair to enhance physical comfort and could enhance performance (Haveman & Kant, 2015; Martins et al., 2014); Room temperature monitoring (Wang, Gu, Ma & Yan, 2017), another factor affecting performance or employee comfort; Monitoring employee health (for example, detecting cough, sneezing, cold, flu, etc.) (Gouma et al., 2017; Costa, Nogueira-Neto & Nohama, 2015; Amoh & Odmae 2015).

Recent articles have criticized human capital analytics (Levenson & Fink, 2017; Boudreau & Wayne, 2017; Minbaeva, 2017), claiming that the many ongoing efforts to analyze human capital by artificial intelligence and machine learning methods have not yielded actionable items, and organizations do not adopt these methods. Among the reasons mentioned are lack of clear methodology, lack of focus, and attempts made by organizations to derive new insights from the data they possess without clearly defining purposes or questions, only to boast using advanced methods such as AI or machine learning. The results are sometimes insufficiently clear and seem to be a "black box", which is why decision makers do not rush to adopt these methods. Even if a certain method proves to be appropriate for a certain department of a certain company, there is no guarantee that it is suitable for other departments. Consequently, reluctance is noticed to adopt methods that have such massive implications without making sure they are better than the existing ones.

In the present paper, we propose a framework that allows focusing on specific issues by modeling the human factor KPIs, and the explanatory factors, which are the KPI components. KPI and explanatory factor measurements, and the digital sources used may differ for each organization or location, depending on the nature and mode of work, and the organizational

culture. The present paper proposes a way to measure digital footprints. Notably, our work does not involve predictions, such as aptness of a candidate for a certain position or team, employee attrition, and course relevance, which are HR practices. We rather analyze the state of the organization in terms of human capital, providing a dynamic (proactive) description of the situation. Our products are similar to those of organizational employee surveys, albeit attained more frequently and with immediate results. They are therefore of interest to the business and link with the organization's goals.

2.2   Human factors

The research of human factors aims to develop a body of knowledge about human attributes, attitudes, abilities and limitations, within a particular context (Chapanis, 1991). As such, it has become a key area in various fields (e.g., psychology, organization and management, engineering), dealing with a variety of topics such as work environment and its design, performance, employee work attitudes, withdrawal behaviors, feedback, leadership, learning and knowledge creation, creativity and innovation. Given this wealth of aspects, individual studies tend to explore a single key aspect within this field. In the present article, we provide an overview of six key human factor constructs – performance, engagement, leadership, workplace relational dynamics, organization developmental support, and learning and knowledge creation – that are particularly relevant to the organizational workplace. This is followed by proposed ways to model these factors (section 3).

A.   Performance

Performance measurement at the individual, group and even organization level is a complex task (Aguinis, 2009). It may take various forms, reflecting numerous perspectives and focal points within performance, such as creativity and service. . A good way to begin conceptualizing performance would be by distinguishing outcomes from behaviors. Recent studies have specifically called to address this point. As Montag et al. (2012) noted (drawing

on performance research by Campbell et al., 1993), "At the heart of this framework is the notion that the workplace criterion space includes both performance behaviors (i.e., behaviors directed toward achieving organizational goals) and outcome effectiveness (i.e., the evaluation of the outcomes of these behaviors along various dimensions such as quantity or quality), two distinct constructs."

We focus on six performance behaviors and outcomes:

1) *Creativity*: "The ability to produce work that is both novel (i.e., original or unexpected) and appropriate (i.e., useful or meets task constraints)" (Sternberg & Lubart, 1991 p. 677). Its key manifestations are the number and originality of ideas;

2) *Innovation*: The implementation of novel ideas (Anderson & West, 1998), namely, their realization in terms of number of new products, revenues derived from newly developed products, and product innovation (incremental, radical);

3) *Service quality*: The extent to which the organization's service staff ensure that customers are satisfied and loyal. It is manifested in the number of repeat purchases, and complaints vs. compliments;

4) *Efficiency*: The ratio of outputs (performed tasks) to inputs (e.g., efforts);

5) *Effectiveness*: Often discussed in terms of goal attainment (Yuchtman & Seashore, 1967);

6) *Organizational citizenship behaviors*: "Behavior(s) of a discretionary nature that are not part of the employee's formal role requirements, but nevertheless promote the effective functioning of the organization" (Organ, 1988, p. 4).

B.   Engagement

The concept of engagement stands for the motivational force underlying a particular activity or work behavior (Kahn, 1990). We include five elements in state engagement:

1) Identification, a state engagement that stands for "the perception of oneness with or belongingness to an organization, where the individual defines him or herself in terms of

the organization(s) in which he or she is a member" (Mael & Ashforth 1992, p.104);

2) Work-family balance, "an overall appraisal of the extent to which individuals' effectiveness and satisfaction in work and family roles are consistent with their life values at a given point in time" (Greenhaus & Allen, 2011, p. 174);

3) Satisfaction, the emotional reaction to the job, by which employees manifest the extent to which they are content with what they do (Locke, 1969);

4) Vitality, the subjective feeling of being alive and alert (Ryan & Frederick, 1997). Vitality may generate a sense of aliveness and energy, denote mental and psychological strength, and result in optimal functioning;

5) Withdrawal intentions (state engagement) "comprise several distinctive yet related constructs (e.g., thinking of quitting, intention to search, and intention to quit), which have been widely studied in relation to withdrawal behavior (e.g., absenteeism, actual turnover)" (Carmeli, 2005, p. 179).

C.   *Leadership*

Scholars have conceptualized the construct of leadership from diverse perspectives. Research of leadership tends to focus on three broadly defined behavioral meta-categories: Task-oriented behaviors, where the primary objective is achieving efficiency and reliability outcomes; relationship-oriented behaviors, where the primary objective is enhancing commitment, trust and cooperation among organizational members; and change-oriented behaviors, where the primary objective is generating major changes that would result in substantial organizational improvements (Yukl, Gordon & Taber, 2002, p. 17). Each meta-category may be manifested by a variety of behaviors. For example, task-oriented leaders focus on tasks that have to be completed and on the implications of outcomes (e.g., efficiency, quality). Similarly, the goals and performance that a leader sets and wishes to pursue determine his task orientation. Relationship-oriented leadership includes feedbacks, namely, leaders' feedbacks may help their

followers to develop and grow. Relationship orientations are cultivated by empowering leaders, who aim to develop, in members, a capability to take the lead in the absence of a formal leader (Manz & Sims, 1987). By supporting such autonomous structures these leaders allow for greater involvement and participation of group members in the decision-making process (Yukl, 1998). Change-oriented leadership articulates a vision that outlines paths defining the organization's identity, strategy and activities. .

Role modeling may be regarded as a meta-construct capable of illustrating task, relationships, and change-oriented behaviors. For example, leaders are role models in displaying task orientation, but they also give clear cues as to how to approach and interact with others (i.e., relationship orientations), and whether they embrace new approaches and ideas and engage in their pursuit.

D.   *Workplace Relational Dynamics*

Relationships are the living tissue that connects members and influences their capacity to thrive in the workplace. Relationships can take many forms: They may be destructive (e.g., contempt) or constructive (e.g., support), depleting or life-giving (Ragins & Dutton, 2007). To illustrate the positive relational dynamics that may emerge and be assessed in the workplace, we focus on four relationship constructs: Trust, psychological safety, connectivity and communication. Trust is defined as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behavior of another" (Rousseau, Sitkin, Burt & Camerer, 1998, p. 395). Members may develop trust in their employer, in leaders and in peers. Psychological safety is the psychological condition that allows people to feel they are safe in taking interpersonal risks and voicing their opinion (Edmondson, 1999). In other words, psychological safety refers to "feeling able to show and employ one's self without fear of negative consequences to self-image, status, or career" (Kahn, 1990, p. 709). Connectivity refers to relationships characterized by openness and generativity (Dutton & Heaphy, 2003).

Connectivity in relationships enables people to see the diverse influences of others as opportunities to learn and grow at work. It involves recognizing the value of relationships in learning new things, generating new ideas, and seeking opportunities to explore and grow (Carmeli & Spreitzer, 2009, p. 174). Communication is probably the most prominent mode of interrelation between people. It is a multidimensional construct, but consists of two specific components: 1) Content of the messages in terms of members' satisfaction with what is being communicated, and 2) Mode by which information is propagate to other members within an organization (Smidts et al., 2001, p. 1052). These features are manifested by the degree to which exchanged and shared information is sufficient, accurate, timely, relevant, and attracts the intended level of attention. There is abundant research evidence suggesting that communication is a key mechanism, by which organizations are able to promote teamwork, job satisfaction and commitment in their members (e.g., Rodwell, Kienzle & Shadur, 1998). Meta-analytical results indicate that communication is related to various dimensions of psychological climate such as cooperation, fairness, and work group warmth (Benzer & Horner, 2015).

E.   *Organizational Support*

Organizational support refers to members' perception of the degree to which an organization appreciates their effort and contribution, and cares about their wellbeing (Eisenberger, Huntington, Hutchison & Sowa, 1986). Organizational support has three facets:

1) Enhancing employee development by a variety of practices such as training, job mobility, and mentoring;

2) Providing instrumental support by allocating the resources, tools and time required to successfully accomplished tasks, while at the same time ensuring people's wellbeing;

3) Behavioral orientation, indicating that the organization values the contribution of individuals, shows interest in their expectations and needs, builds their confidence, and gives them a sense of ownership.

F.   Learning and Knowledge Creation

Learning is a process whereby new knowledge is created, exchanged and integrated (Argote, 1999). Learning may take different forms, such as learning from failure vs. learning from success, or learning from direct experience vs. learning from indirect experience. Each mode of learning implies different processes. Three fundamentals define the knowledge creation process: access to knowledge, exchange of knowledge, and the combination of exchanged knowledge (Nonaka & Takeuchi, 1995; Smith, Collins & Clark, 2005). To determine who knows what and how reliable the knowledge is, knowledge bases may have to be unraveled.

3.   The Proposed Ontology

Utilizing a commonly used methodology, as specified above (see also Figure 1), we have put together an ontology of human resource factors, in an attempt to advance workplace analytics research. Figure 1 shows the components we used to identify the key performance indicators (KPIs) of human capital. The first step of the modeling process was conceptually defining each human capital KPI. Next, we identified key components that underlie the KPIs of human capital. For example, the key components underlying organizational support are support directed at employee development, means (instrumental) support and behavioral support. The last component comprises digital sources that yield the data required to assess latent variables, and their components and manifestations. Although the present ontology is rather complex (i.e., comprised of six main components that are further broken down into three sub-level elements), we made an effort to present it in a simplified way. To this end, we focused on one KPI – Engagement (Table 3), on two explanatory factors of engagement (satisfaction and vitality), and within them on manifestations of emotional response to work, as proxied by email sentiments. These form two pathways of email sentiment: 1) Raw source of email correspondences (digital footprint) → emotional response (manifestation) → satisfaction (explanatory factor) → engagement (KPI); 2) Raw source of email correspondences (digital

footprint) → sense of aliveness (manifestation) → vitality (explanatory factor) → engagement (KPI).

Research studies frequently use surveys to assess human factors. We propose a different approach that is based on assessing the digital footprints of various human factors. Tables 1 to 6 present ontology of six human factors: performance, engagement, leadership, workplace relational dynamics, organizational developmental support, and learning and knowledge creation. Drawing on the existing literature, we list the explanatory factors for each KPI and their manifestations. They are shown in the upper block of Figure 1, where the KPI is broken down into digital factors, each depicted by specific manifestations. A potential digital source is then recommended to assess each factor. The digital sources refer to: Calendar, eMail, Forums & Portals, HR reports; Manuals, Quality Assurance, Releases, Tasks (PMO), and others (noted as "z"). The frequency of the data collection use the notation: Y =Year, Q =Quarter, M =Month, W =Week. The formatting representation may be: Abs =Absolute values, Delta = Absolute Changes, Per = Percentage Change, etc.

----- Insert Tables 1 to 6 about here -------

Tables 1 to 6 map out the six proposed KPIs. Table 1 illustrates the performance factor. For example in assessing human performance, we may evaluate members' creativity, innovation, service quality, efficiency and effectiveness, and extra-role behaviors (citizenship). To assess creativity we may use managers' evaluation of their employees. This can be done on an annual basis, but in organizations where performance evaluation is carried out more frequently (e.g., every six months) assessment may be adjusted accordingly. To assess innovation, we may use organizational and external records to tabulate the number of newly developed products/services, the number of patents by IP submissions, product quality by QA reports, sales of new products using the financial statement, and development speed by the PMO. Service quality may be assessed by probing the level of customer loyalty and satisfaction, and

the delta in sales relative to services. Customer loyalty may be assessed based on meeting cancellations and fluctuation or migration (churn) on a periodic basis. Customer satisfaction may be assessed by organizational records that contain digital service evaluation forms and complimentary letters. Change in sales may be assessed annually using the financial reports. To asses efficiency, the following statistics may be used: Completion of tasks on time based on PMO data, use of allocated resources to complete tasks, based on PMO budgetary evidence, quality of completed tasks based on QA reports, and task load based on number of weekly overtime hours. Effectiveness may be assessed by evaluation of tasks completed on time within or beyond the allocated budget. Finally, organizational citizenship behaviors may be assessed by peer behavior evaluations and members' responses to queries in organizational portal forums.

4.   Applying Sentiment analysis to Organizational eMails

4.1   Text mining for sentiment analysis

Sentiment analysis aims to identify the emotional attitude of text authors regarding certain phenomena, such as movies they watch or events occurring in organizations for which they work (Pang, Lee & Vaithyanathan, 2002; Ghazizadeh, McDonald & Lee, 2014). Towards developing our model we analyzed employee emails from the Enron email corpus (Klimt & Yang, 2004), to identify the general sentiment of the body of Enron employees (and not of individual employees). To this end, we assessed the sentiment of each individual email over a few weeks, and aggregated them to create a collective picture.

A widely used approach in building sentiment analysis solutions is that of using *classifiers* – a machine learning tool (Liu & Zhang, 2012). Classifiers may be trained using a *loss function* – a function that expresses an error of the classifier over the training data. In this case, classifier training simply seeks to minimize the loss function. In the case of a Support Vector Machine classifier, the loss function is smooth and differentiable, and may be minimized as described in

Gu & Sheng 2017, or in Gu, Sheng, Tay & Li, 2015. However, when the search for a solution takes the form of a combinatorial optimization problem, other families of optimization algorithms are used. One effective approach is based on combining a Genetic Algorithm and Ant Colony Adaptive Collaborative Optimization (MGACACO) algorithm (Deng, Zhao, Zou, Li, Yang & Wu 2017), used to solve the traveling salesman problem. Another approach uses an improved adaptive Particle Swarm Optimization (DOADAPO) algorithm that is applied to the airport gate assignment problem (Deng, Zhao, Yang, Xiong, Sun & Li 2017). Yet another effective way to solve global optimization problems is developed in Xue, Jiang, Zhao & Ma 2017, where self-adaptive techniques are used to improve the Artificial Bee Colony algorithm with an application to a real clustering problem, based on *K-means* technique.

To tune their parameters, classifiers require a *tagged corpus* as input, namely, a body of text documents, each of which is given a single *sentiment tag* (e.g. positive, negative or neutral) (Calix, Javadpour & Knapp 2011). In constructing a sentiment analysis for the Enron corpus, we were challenged by its lack of balance: only a very small percentage of the emails could be categorically tagged as displaying a positive or negative sentiment, whereas the vast majority was neutral. The rate of non-neutral emails was below 0.1%. As a rule, it is very hard to achieve high classification accuracy from unbalanced corpora such as this (Chawla, 2010).

Another challenge was that no email corpus was available that was manually tagged for sentiment. To overcome this problem, we used a tagged corpus from another domain that was open to the public – sentiment analysis in movie reviews (Cornell Movie Review Data, 2012). Our choice was based on the observation that the *sentiment lexicon* – the set of words used to describe sentiments – is largely domain-independent. However, when applied to the Enron email corpus, our sentiment analysis as trained on the movie corpus proved inaccurate. We therefore decided to adjust the movie corpus by adding to it a small subset from the email corpus. We manually tagged the added emails and used the adjusted corpus to construct a new

classifier. The accuracy of the new classifier proved much better: about half of the emails tagged as displaying a positive or a negative sentiment, corresponded to their sentiment tags, whereas the vast majority of the neutral emails had indeed been tagged as *neutral*. This level of accuracy was sufficient to reveal fluctuations in sentiment among the Enron employee body over time (e.g., after the CEO was replaced, or once the employees had accepted the new CEO). We used the correlation between our sentiment analysis over time and important events in Enron's history to validate our approach of estimating overall employee sentiment in an organization.

4.2 The text mining algorithm

To build an organizational sentiment analysis model, we used a naïve Bayes classifier (Zhang 2004). Let c be the set of classes. The naïve Bayes classifier regards each document as the set of all its words. It also assumes that for each word w, the probability of observing w in document d, given class c, may be written as follows:

$$\Pr(d \mid c) = \prod_{w \in d} \Pr(w \mid c) \quad (1)$$

This assumption means that given class c, words in the document are independent of other words in the document, their relative position in d, the length of the document and any other context of the document. This assumed independency gave rise to the name of the classifier.

Next, based on the Bayes theorem (Mitchell 1997), it is possible to write the probability of a class in a given document, as follows:

$$\Pr(c \mid d) = \frac{\Pr(c) \Pr(d \mid c)}{\Pr(d)} \quad (2)$$

Using this expression, the definition of the classification function is: given document d, choose class c that maximizes the above probability

$$classify(d) = \max_c \Pr(c \mid d) \quad (3)$$

Since the denominator of expression (2) does not depend on c, the classification function may be rewritten as follows

$$classify(d) = \max_c \Pr(c) \Pr(d \mid c) \quad (4)$$

Using (1) we get

$$classify(d) = \max_c \Pr(c) \prod_{w \in d} \Pr(w \mid c) \quad (5)$$

To make the above derivations applicable in practice, the probabilities $\Pr(c)$ and $\Pr(w|c)$ are estimated based on a *training corpus* that consists of a set of documents D, where each document $d \in D$ is assigned a class $c \in C$.

4.3 The Enron email corpus

As mentioned, we used the Enron email corpus to validate our argument that human capital KPIs can be evaluated by analyzing information available in the organization's systems. The Enron email dataset was made public by the Federal Energy Regulatory Commission while the company was under investigation. Originally, the database had comprised over 600,000 diverse emails, both personal and official, generated by 158 employees. Some of the emails were deleted in a redaction effort prompted by employee requests (for a more comprehensive review of this dataset see Keila & Skillicorn, 2005; Sebastian, Bródka & Kazienko, 2011; Shetty & Adibi, 2004). In our analysis, we used a clean version of the dataset containing 250,000 email messages generated by 151 employees.

First we had to isolate a unit in the organization and analyze its e-mail correspondence. By examining the overall body of e-mail correspondence, we identified employee groups who exchanged messages frequently. A closer look at the group's messages enabled us to identify the manager. By further examining the manager's focal group messages, we were able to verify that we indeed identified the entire unit. We used external web data (e.g., LinkedIn) to corroborate our findings. At the end of the process we succeeded in extracting the organization's legal unit and identifying its manager.

Our second task was identifying significant events that occurred over the period covered by the corpus, namely, week 32 of 2000 to week 9 of 2002.

The events we chose to relate to were:

      A: W:7 Y:2001 – Skilling appointed Enron CEO.

      B: W:33 Y:2001 – Skilling resigns, Lay reappointed CEO.

      C: W:42 Y:2001 – Inquiry launches.

      D: W:48 Y:2001 – Enron goes bankrupt, thousands of workers laid off.

5.   Results - Vitality and Satisfaction as Manifested in Enron's Email

Towards achieving the desired results, we conducted the following experiments:

First, we looked to evaluate *vitality* (V), which is an explanatory factor of engagement. We used

the ratio between emails sent during off-work and work hours; i.e. $V = \dfrac{E_0}{E_W}$

Figure 2 shows vitality to be an explanatory factor for the group and the manager, indicating that vitality was significantly affected by the events. Prior to Skilling's resignation and the Securities Commission inquiry, significant changes had occurred in employee vitality. An interesting observation concerns the manager's level of vitality, which had changed about two weeks earlier than that of his/her group. This may indicate that the manager had access to information on the state and functioning of the organization that was unavailable to other employees.

------   Insert Figure 2 about here   ------

The second experiment involved *sentiment* in the company. It analyzed the aggregated sentiments that came up in the emails, and studied them against the events listed above. Figure 3, which measures satisfaction, shows how the events affected sentiment in the company. As mentioned, event A denotes the appointment of a new CEO. In situations where CEOs are dismissed due to major problems, one might expect uncertainty accompanied by a certain amount of hope for change and ultimately for better outcomes. As Figure 3 shows, negative sentiment increased in the organization during the adaptation period, but then gradually subsided and went below the starting point, implying that the new CEO was accepted as capable

of leading the company in a better direction. However, once the CEO resigned due to increased pressure on the company, the reappointment of the former CEO (Event B) was not as successful, as members seemed to have lost hope, and were distrustful of this change. As the results indicate, sentiment never recovered before Enron finally crashed.

------   Insert Figure 3 about here   ------

6.   Discussion and Conclusion

Schneider's claim (1987) that "the people make the place" remains as true as it was several decades ago. Understanding people's attitudes, intentions, and behaviors is therefore fundamental to cultivating better work processes and outcomes. People's perceptions about their work and organization shape their behavior in the workplace, which in turn has implications for events that occur in their units and organizations, and for their functioning. Yet interpreting people's perceptions and behaviors is a complex task. Conventional tools such as survey-based data collection to assess employee perceptions require substantial resources. What is more, they have limitations that call for caution in interpreting the data, since the information, being subjective, is often inflated and biased, and real-time assessment is seldom feasible. To address this issue, large companies such as Google® and IBM® have been developing data mining procedures, intended to equip organizations with less costly and more reliable tools that would enhance their understanding of their human resources and help manage them in a way that would improve the outcomes. Following this line of thought, the new approach of the present study proposes that organizations and researchers use data mining techniques (sentiment analysis and opinion mining) that would enable them to trace emergent patterns and evaluate changes in various human factors. This is done by modeling key performance indicators and defining their explanatory factors, manifestations and diverse corresponding digital footprints. We used the Enron email corpus as a case-in-point to demonstrate the feasibility of our approach by showing how digital footprints can serve to trace

satisfaction and vitality, which are explanatory factors of the engagement KPI. To determine employee satisfaction, we identified negative sentiment levels by period, indicating low satisfaction. To determine vitality, we analyzed the change in the ratio between off-work emails to working hour emails and identified changes in the employees' vitality against critical company events. Although the feasibility test run on the Enron email corpus demonstrated the predictability of two manifestations relevant to the engagement KPI, and was limited to a single digital footprint (e-mail correspondence), it confirmed that our method was potentially useful in understanding and analyzing human factors within organizations.

Our approach offers a reliable and convenient way to evaluate human factors by using digital sources and footprints available in any organization's information systems. Furthermore, it also enables an integrated view of numerous perspectives indicating levels of individual and group behaviors in organizations. By adopting this method, an organization enhances its capacity of tracing and predicting emerging behavioral patterns. This, in turn, enables the organization to engage in "preventive actions" or "promoting actions" that are capable molding behaviors towards a desired end. For example, tracking the way organization members come to accept a new management team could suggest what kind of messages should be communicated to the employees, to persuade them that the new strategic orientation is robust, and elicit further engagement in the new direction.

While our approach and method offer some obvious advantages, one should also be aware of potential limitations or hurdles that could be encountered in their implementation. One critical issue regards members' privacy. Concerns for privacy are liable to preempt the willingness and ability to implement the full model, thus impeding the full realization of our method's potential. We must therefore guarantee that the privacy of individuals and groups is meticulously preserved, by taking one or several of the following measures. First, the data should be extracted and analyzed in a way that would not expose individual identities, namely, data should carry

only technical identifiers and all personal identification must be deleted (see also Yuan, Sun & Lv, 2016). Second, the data should be aggregated and avoid any reference to individual members Third, the analysis and presentation of the results should focus on patterns of behaviors rather than on exact numerical values. Fourth, instead of content-based processing, it is recommended that organizations and researchers adopt technical text-based processing, similar to anti-virus or fraud detection programs that search for patterns in the text rather than explore its content.

Further research is required to substantiate the assumption that digital data analysis provides equally reliable results as traditional subjective survey reports. While we conducted a test of principle that showed feasibility for one specific element, future research should further explore this method. Scholars in a variety of fields have used proxies to evaluate socio-psychological concepts (e.g., "rewards are viewed as proxies for goals/targets and outcomes", see Samnani & Singh, 2014). However, these proxies should be employed with caution as a systematic analysis is required, to assess the relationship between a proxy and the behavior it aims to capture.

Three key avenues should be explored in future research. The first involves a systematic examination of all theoretical factors to support each with a corresponding model. Second, each model should be examined in different environmental settings such as knowledge-intensive organizations, public sector organizations, and non-governmental organizations. Finally, for a more robust evaluation of the power of our approach and method, the model findings should be examined against annual surveys conducted in different organizations.

###

Conflicts of interest: None

References

Aguinis, H. (2009). *Performance management*. Upper Saddle River, NJ: Pearson/Prentice Hall.

Amoh, J., & Odame, K. (2015, October). DeepCough: A deep convolutional neural network in a wearable cough detection system. In *Biomedical Circuits and Systems Conference (BioCAS), 2015 IEEE* (pp. 1-4). IEEE.

Anderson, N., & West, M. A. (1998). Measuring climate for work group innovation: development and validation of the team climate inventory. *Journal of Organizational Behavior, 19*, 235–258.

Ara, K., Akitomi, T., Sato, N., Tsuji, S., Hayakawa, M., Wakisaka, Y., ... & Yano, K. (2011, January). Healthcare of an organization: Using wearable sensors and feedback system for energizing workers. In *Proceedings of the 16th Asia and South Pacific Design Automation Conference* (pp. 567-572). IEEE Press

Argote, L. (1999). *Organizational Learning: Creating, Retaining and Transferring Knowledge*. Boston: Kluwer Academic.

Benzer, J., &Horner, M. (2015). A Meta-analytic integration and test of psychological climate dimensionality. *Human Resource Management, 54*, 457–482.

Bollinger, J., Hardtke, D., & Martin, B. (2012, October). Using social data for resume job matching. In *Proceedings of the 2012 workshop on Data-driven user behavioral modelling and mining from social media* (pp. 27-30). ACM.

Boudreau, J., & Cascio, W. (2017). Human capital analytics: why are we not there?. *Journal of Organizational Effectiveness: People and Performance*, *4*(2).

Brutus, S., Gill, H., & Duniewicz, K. (2010). Self-reported limitations in industrial and organizational psychology. *Personnel Psychology, 63*, 907-936.

Brutus, S., Aguinis, H., & Wassmer, U. (2013). Self-reported limitations and future directions in scholarly reports: Analysis and recommendations. *Journal of Management, 39*(1), 48-75.

Calix, R, A., Javadpour, L., Knapp, G. M. (2011). Detection of affective states from text and speech for real-time human–computer interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 54*, 530

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco: Jossey-Bass.

Carmeli, A. (2005). The relationship between organisational culture and withdrawal intentions and behavior. *International Journal of Manpower, 26*(2), 177-195.

Carmeli, A., & Spreitzer, G. M. (2009). Trust, connectivity, and thriving: Implications for innovative behaviors at work. *Journal of Creative Behavior, 43* (3), 169-191.

Chan, C. F., & Eric, W. M. (2010, August). An abnormal sound detection and classification system for surveillance applications. In *Signal Processing Conference, 2010 18th European* (pp. 1851-1855). IEEE.

Chapanis, A. (1991). To communicate the human factors message, you have to know what the message is and how to communicate it. *Bulletin of the Human Factors Society, 34*, 1-4.

Costa, A., & Veloso, A. (2015). Employee Analytics through Sentiment Analysis. In *SBBD* (pp. 101-112).

Costa, T. D., Nogueira-Neto, G. N., & Nohama, P. (2015, August). Cough detection through mechanomyographic signal in synchronized respiratory electrical stimulation systems. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (pp. 4590-4593). IEEE

Davenport, T. H., Harris, J. & Shapiro, J. (2010). Competing on talent analytics. *Harvard Business Review, 88*, 52–58.

Deng, W., Zhao, H., Yang, X., Xiong, J., Sun, M., & Li, B. (2017). Study on an improved adaptive PSO algorithm for solving multi-objective gate assignment. *Applied Soft Computing, 59*, 288–302.

Deng, W., Zhao, H., Zou, L., Li, G., Yang, X., & Wu, D. (2017). A novel collaborative optimization algorithm in solving complex optimization problems. *Soft Computing*, *21*(15), 4387-4398.

Diego-Mas, J. A., & Alcaide-Marzal, J. (2014). Using Kinect™ sensor in observational methods for assessing postures at work. *Applied ergonomics*, *45*(4), 976-985.

Dutton, J. E., & Heaphy, E. D. (2003). The power of high-quality connections at work. In K.S. Cameron, J. E. Dutton, & R. E. Quinn (Eds.), *Positive organizational scholarship* (pp. 263-278). San Francisco: Berrett-Koehler Publishers.

Edmondson, A.C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly, 44* (2), 350-383.

Eidinger, E., Enbar, R., & Hassner, T. (2014). Age and gender estimation of unfiltered faces. *IEEE Transactions on Information Forensics and Security*, *9*(12), 2170-2179.

Eisenberger, R., Huntington, R., Hutchison, S., & Sowa, D. (1986). Perceived organizational support. *Journal of Applied Psychology, 71*, 500–507.

Feloni R. (Jun. 28, 2017), Consumer-goods giant Unilever has been hiring employees using brain games and artificial intelligence - and it's a huge success, Business Insider, http://www.businessinsider.com/unilever-artificial-intelligence-hiring-process-2017-6

Gallup, Inc. (2013). *State of the American Workplace: Employee Engagement Insights for U.S. Business Leaders*. Retrieved January 20, 2014 from: http://www.gallup.com/strategicconsulting/163007/state-american-workplace.aspx

Gartner, Workforce Analytics, https://www.gartner.com/it-glossary/workforce-analytics

Gelbard, R., & Carmeli, A. (2008). Toward an Ontology of ICT Management: Integration of Organizational Theories and ICT Core Constructs. In P. Rittgen (Ed.), *Handbook of ontologies for business interaction* (pp. 157-171). IGI Global.

Geng, X., Zhou, Z. H., & Smith-Miles, K. (2007). Automatic age estimation based on facial aging patterns. *IEEE Transactions on pattern analysis and machine intelligence*, *29*(12), 2234-2240.

Ghazizadeh, M., McDonald , A. D., Lee, J. D. (2014). Text mining to decipher free-response consumer complaints: Insights from the NHTSA vehicle owner's complaint database. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 56*, 1189

Gouma, P. I., Wang, L., Simon, S. R., & Stanacevic, M. (2017). Novel Isoprene Sensor for a Flu Virus Breath Monitor. *Sensors*, *17*(1), 199.

Greenhaus, J.H., & Allen, T.D. (2011). Work-family balance: A review and extension of the literature. In Campell Quick J, Tetrick LE (Eds.), *Handbook of Occupational Health Psychology* (pp. 165–183). Washington, DC: American Psychological Association.

Gu, B., & Sheng, V. S. (2017). A robust regularization path algorithm for v-support vector classification. *IEEE Transactions on neural networks and learning systems*, *28*(5), 1241-1248.

Gu, B., Sheng, V. S., Tay, K. Y., Romano, W., & Li, S. (2015). Incremental support vector learning for ordinal regression. *IEEE Transactions on Neural networks and learning systems*, *26*(7), 1403-1416.

Haveman, S., & Kant, G. (2015). Smart monitoring of worker posture in an office environment. *viewed on Jul*, *31*, 1-17.

HireVue - video interviewing, assessments, and coaching   https://www.hirevue.com/

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal, 33*, 692-724.

Keila P.S. & Skillicorn D.B., (2005). "Structure in the Enron Email Dataset". Computational & Mathematical Organization Theory, 11(3), 183-199.

Klimt B. and Yang Y., (2004). Introducing the Enron Corpus. CEAS. Retrieved from: http://www.**ceas**.cc/papers-2004/168.pdf

Koriat, N., & Gelbard, R. (2014). Knowledge sharing motivation among IT personnel: Integrated model and implications of employment contracts. *International Journal of Information Management*, 34(5), 577-591.

Koriat, N., & Gelbard, R. (2017). Knowledge Sharing Analytics: The Case of IT Workers. *Journal of Computer Information Systems*, 1-11.

Leftheriotis, I., & Giannakos, M. N. (2014). Using social media for work: Losing your time or improving your work?. *Computers in Human Behavior*, *31*, 134-142.

Levenson, A., Levenson, A., Fink, A., & Fink, A. (2017). Human capital analytics: too much data and analysis, not enough models and business insights. *Journal of Organizational Effectiveness: People and Performance*, *4*(2), 145-156.

Locke, E.A. (1969). What is job satisfaction? *Organizational Behavior and Human Performance. 4*, 309-336.

Mael, F., & Ashforth, B. E. (1992). Alumni and their alma mater: a partial test of the reformulated model of organizational identification. *Journal of Organizational Behavior, 13*, 103–123.

Manz, C. C., & Sims, H. P., Jr. (1987). Leading workers to lead themselves: The external leadership of self-managing teams. *Administrative Science Quarterly, 32*, 106−128.

Martins, L., Lucena, R., Belo, J., Almeida, R., Quaresma, C., Jesus, A. P., & Vieira, P. (2014). Intelligent Chair Sensor–Classification and Correction of Sitting Posture. In *XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013* (pp. 1489-1492). Springer, Cham.

Mattmann, C., Amft, O., Harms, H., Troster, G., & Clemens, F. (2007, October). Recognizing upper body postures using textile strain sensors. In *Wearable Computers, 2007 11th IEEE International Symposium on* (pp. 29-36). IEEE

Minbaeva, D., & Minbaeva, D. (2017). Human capital analytics: why aren't we there? Introduction to the special issue. *Journal of Organizational Effectiveness: People and Performance*, *4*(2), 110-118.

Mitchell T. (1997). *Machine Learning*. New York: McGraw-Hill.

Moniz, A., & Jong, F. (2014). Sentiment analysis and the impact of employee satisfaction on firm earnings.

Montag, T., Maertz, C.P., Jr., & Baer, M. (2012). A critical analysis of the workplace creativity criterion space. *Journal of Management, 38*, 1362-1386.

Mota, S., & Picard, R. W. (2003, June). Automated posture analysis for detecting learner's interest level. In *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Conference on* (Vol. 5, pp. 49-49). IEEE.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. New-York: Oxford University Press.

Organ, D. W. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington Books.

Pang B., Lee L., & Vaithyanathan, S., (2002). Thumbs up? Sentiment classification using machine learning techniques. *Proceedings of EMNLP*, 79-86.

Piazza F., & Strohmeier S., (2013). Domain-driven data mining in human resource management: A review. *Expert Systems with Applications, 40*(7), 2410-2420.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology, 63*(1), 539-569.

Preece, S. J., Goulermas, J. Y., Kenney, L. P., Howard, D., Meijer, K., & Crompton, R. (2009). Activity identification using body-mounted sensors—a review of classification techniques. *Physiological measurement*, *30*(4), R1.

Punnoose, R., & Ajit, P. (2016). Prediction of employee turnover in organizations using machine learning algorithms. *International Journal of Advanced Research in Artificial Intelligence (ijarai)*, *5*(9).

Quwaider, M., & Biswas, S. (2008, March). Body posture identification using hidden Markov model with a wearable sensor network. In *Proceedings of the ICST 3rd international conference on Body area networks* (p. 19). ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

R. The R Project for Statistical Computing. http://www.r-project.org/

Ragins, B. R., & Dutton, J. E. (2007). Positive relationships at work: An introduction and invitation. In J. E. Dutton & B. R. Ragins, Eds. *Exploring positive relationships at work: Building a theoretical and research foundation* (pp. 1-25). Mahwah, NJ: Lawrence Erlbaum.

Robinson, S. D., Sinar, E., & Winter, J. (2014). Social Media as a Tool for Research: A Turnover Application Using LinkedIn. *TIP: The Industrial-Organizational Psychologist*, *52*(1), 133-141.

Rodwell, J. J., Kienzle, R., & Shadur, M. A. (1998). The relationship among work-related perceptions, employee attitudes, and employee performance: The integral role of communications. *Human Resource Management, 37*, 277–293.

Rosario, J. L., Diógenes, M. S. B., Mattei, R., & Leite, J. R. (2014). Differences and similarities in postural alterations caused by sadness and depression. *Journal of bodywork and movement therapies*, *18*(4), 540-544

Rosário, J. L., Diógenes, M. S. B., Mattei, R., & Leite, J. R. (2016). Angry posture. *Journal of bodywork and movement therapies*, *20*(3), 457-460

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review, 23*, 393-404.

Ryan, R.M., & Frederick, C. (1997).On energy, personality and health: Subjective vitality as a dynamic reflection of well-being. *Journal of Personality, 65*, 529–565.

Samnani, A. K., & Singh, P. (2014). Performance-enhancing compensation practices and employee productivity: The role of workplace bullying. *Human Resource Management Review, 24*(1), 5-16.

SAP HANA Predictive Analytics Library (2014). http://help.sap.com/hana

Schneider, B. (1987). The people make the place. *Personnel Psychology, 40*, 437-453.

Sebastian Palus S., Bródka P., & Kazienko P., (2011). Evaluation of organization structure based on email interactions. *International Journal of Knowledge Society Research, 2*(1), 1-13.

Shetty J., & Adibi J. (2004). The Enron email database schema and brief statistical report. *Information Sciences Institute Technical Report*, University of Southern California.

Smidts, A., Pruyn, A. H., & van Riel, C. B. M. (2001). The impact of employee communication and perceived external prestige on organizational identification. *Academy of Management Journal, 44*, 1051–1063.

Smith, K.G., Collins, C.J., & Clark, K.D. (2005). Existing knowledge, knowledge creation capability and the rate of new product introduction in high technology firms. *Academy of Management Journal, 48*(2), 346-357.

Sternberg, R.J., & Lubart, T. I. (1991). An investment theory of creativity and its development. *Human Development, 34*, 1-31.

Subhashini, R., & Niveditha, P. R. (2015). Analyzing and Detecting Employee's Emotion for Amelioration of Organizations. *Procedia Computer Science*, *48*, 530-536

Sullivan, J. (2013). How Google is using people analytics to completely reinvent HR. *TLNT*,

Feb 26, 2013. Retrieved January 20, 2014 from: http://www.tlnt.com/2013/02/26/how-google-is-using-people-analytics-to-completely-reinvent-hr/

Tian, Q., & Chen, S. (2017). Cross-heterogeneous-database age estimation through correlation representation learning. *Neurocomputing*, *238*, 286-295.

Wang, B., Gu, X., Ma, L., & Yan, S. (2017). Temperature error correction based on BP neural network in meteorological wireless sensor network. *International Journal of Sensor Networks*, *23*(4), 265-278.

Xue, Y., Jiang, J., Zhao, B., & Ma, T. (2017). A self-adaptive artificial bee colony algorithm based on global best for global optimization. *Soft Computing*, 1-18.

Yi, X., Allan, J., & Croft, W. B. (2007, July). Matching resumes and jobs based on relevance models. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 809-810). ACM.

Yuan, C., Sun, X., & Lv, R. (2016). Fingerprint liveness detection based on multi-scale LPQ and PCA. *China Communications*, *13*(7), 60-65.

Yuchtman, E., & Seashore, S. E. (1967). A system resource approach to organizational effectiveness. *American Sociological Review, 32*, 891-903.

Yukl, G. (1998). *Leadership in organizations* (4th ed.). Upper Saddle River, NJ: Prentice-Hall.

Yukl, G., Gordon, A., & Taber, T. (2002). A hierarchical taxonomy of leadership behavior: Integrating a half century of behavior research. *Journal of Leadership & Organizational Studies, 9*(1), 15-32.

Zhang H. (2004). The optimality of naive bayes. *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference FLAIRS*. AAAI Press.

van Zoonen, W., Verhoeven, J. W., & Vliegenthart, R. (2016). How employees use Twitter to talk about work: A typology of work-related tweets. *Computers in Human Behavior*, *55*, 329-339.