THE PATTERNS OF ASSOCIATION BETWEEN STRATEGIC ORIENTATIONS, INNOVATION, AND FIRM PERFORMANCE: A CO-OCCURRENCE NETWORK ANALYSIS APPLICATION

Abstract

Strategic orientation represents market-based resources embedded into organizational culture which can be examined through company´s annual reports. This paper aims to identify the relevance and association patterns of strategic orientations, innovation and firm performance to better understand how these constructs are adopted and exhibited in practice. Drawing on annual reports of 48 firms within S&P 500's, a co-occurrence network analysis is conducted to statistically and visually extract information from the text data. Results indicate that market orientation plays a central role in the associations, linking other orientations with innovation and firm performance. Unexpectedly, innovation is not closely associated with performance. Three major co-occurrence patterns of association were identified, suggesting that customers, competitors, and resources within firms cover a central place in corporate narratives. Implications for practitioners and researchers are discussed and future lines of research are presented.

***Keywords*: strategic orientation, innovation, text mining.**

INTRODUCTION

The concept of strategic orientation (SO) has attracted widespread attention from market, entrepreneurial and management scholars (Hakala, 2011). By definition, strategic orientations, namely, market, entrepreneurial and learning orientations are linked to firm performance as its antecedents and important drivers (Hult et al., 2004). Briefly, market orientation (MO) entails a culture and behaviors in which customers, competitors and markets are the center of a firm's activity (Deshpande & Webster, 1989; Gnizy et al., 2014); entrepreneurial orientation (EO) involves the propensity to pursue new market opportunities (Covin & Slevin, 1991; Lumpkin & Dess, 1996); and learning orientation (LO) has to do with the inclination to create and use knowledge (Calantone et al., 2002; Sinkula et al., 1997). Innovation, as an outcome, comprises the consequences of innovation activities or the outputs of innovation process (Crossan & Apaydin, 2010), whereas firm performance captures the underlying manifestations of how well a firm is effectively satisfying its stated goals (Bergh et al., 2016; Combs et al., 2005).

The literature on SO has established the notion that market, entrepreneurial and learning orientations are multidimensional, interlinked, correlated, but distinct constructs (Baker & Sinkula, 2009; Hakala, 2011), and share similar characteristics, in terms of their role on innovation (Gatignon et al., 2016). The association between SO and firm performance might be stronger when they are considered collectively rather than in isolation and in an operating interplay basis (e.g., Cambra-Fierro et al., 2012; Ho et al., 2015; Mu & di Benedetto, 2011).

Past research has stated that when strategic orientations are operating synergistically, innovation could benefit from complementarity, which means that the effect of one orientation can increase the effectiveness and efficiency of other orientations and that the combination of strategic orientations leads to superior performance (e.g., (Baker & Sinkula, 2009; Gnizy & Shoham, 2014; Ho et al., 2015; Mu & di Benedetto, 2011).

Despite the growing existence of empirical research, little is known about the interrelationships between SO (Grinstein, 2008; Hakala, 2011), and innovation linking SO with firm performance. In this sense, the identification of complementarities among the various orientations is critical for examining their synergies (Gatignon et al., 2016). Particularly, further research is requested on exploring how the more successful firms adopt and balance various combinations of strategic orientations to develop a more complex corporate culture (Grinstein, 2008).

Following Short et al. (2009) and Pollach (2012), this research draws from the assumption that the presence of a strategic orientation in a firm should be highlighted in its corporate disclosures as a reflection of its managerial cognitions, organizational culture, values, or identity. In other words, strategic orientations can be examined through their projections or exhibitions in the key organizational narratives –e.g., annual reports–. These key narratives are a source of rich and valuable data, from a qualitative point of view, which surveys or interviews cannot provide in the same manner. One source of valuable unstructured data is the company's annual report on Form 10-K, which offers a detailed picture of a company's business and discusses its perspective on the business results and what is driving them (U.S. Securities and Exchange Commision, 2019).

The availability of this kind of rich data sources combined with the growing sophistication of analytical techniques due to recent advances in both the computational power and mathematical models and algorithms for the collection, extraction, visualization, analysis and interpretation of data (Castelfranchi, 2017; Filippov & Hofheinz, 2016) provide to the researchers the opportunity to explore and test hypotheses in new contexts and gain valuable insights that were difficult to attain with more traditional research methods (Duriau et al., 2007). The application of new or more advanced CATA methods offer new options for organizational researchers to uncover the latent themes and associations in a body of text (Short et al., 2018).

Text mining, defined as the "discovery and extraction of interesting, non-trivial knowledge from free or unstructured text" (Kobayashi et al., 2018), is a possible way to produce knowledge derived from textual patterns and relationships, and can be used to reveal facts, trends, or constructs (Delen & Crossland, 2008; Kobayashi et al., 2018). Results derived from text mining applications are data-driven and not researcher-driven, improving the transparency of the evidence used to support research conclusions (Pokorny et al., 2018).

This exploratory data-driven paper aims to address these research gaps by identifying the relevance and centrality of SO, innovation and firm performance and extracting their co-occurrence patterns, in order to better understand how these constructs of interest are adopted, combined and balanced in business practice.

Drawing on Form 10-K annual reports of 48 firms within S&P 500's communication services and materials industry sectors, this paper conducts a co-occurrence network analysis using text analysis software KH Coder (Higuchi, 2017) to statistically and visually extract information from the text data. Specifically, this research tackles the following research questions: To what extent firms adopt and exhibit their strategic orientation, innovation and firm performance? Which patterns of strategic orientations are adopted by firms and how these patterns are associated with innovation and firm performance?

THEORETICAL FRAMEWORK

Under the resource-based theory of the firm (RBT), the strategic orientations (SO) represents important market-based resources –assets or capabilities– (Hult & Ketchen, 2001; Kozlenkova et al., 2014; Lonial & Carter, 2015) embedded into an organizational culture reflecting "the strategic directions implemented by a firm to create the proper behaviors for the continuous superior performance of the business" (Gatignon & Xuereb, 1997); and "the means by which firms choose to attempt to create a sustainable presence in the markets in which they compete" (Gnizy et al., 2014). SO reflects elements that influence the ability to develop and market innovations of a firm more effectively, and these innovation outcomes in turn lead to greater overall firm performance. Within the RBT framework, SO have two key characteristics: 1) intangibility, and 2) complementarity.

First, as intangible resources, SO cover an enhanced set of skills deeply ingrained into the everyday routines of an organization which are difficult to imitate by competitors (Zhou et al., 2005). Thus, organizations possessing and deploying strategic orientations should have sustained competitive advantages over competitors that lack such resources (Barney, 1991). In this sense, strategic orientations offer great potential to develop into competitive advantage and superior firm performance.

Second, SO are complementary. RBT explains how SO combine to create a competitive advantage for a firm under the notion of bundling resources. Individual strategic orientations are necessary, but insufficient, conditions for business success (Hult & Ketchen, 2001). This study is in line with the RBT's market-based resources characteristics.

Few existing studies simultaneously and in a complementary perspective analyzed MO, EO, and LO (Deutscher et al., 2016). SO can exist and support each other at the same time in one organization (Leenders et al., 2016). Organizations combining several orientations perform better than those focusing on a single orientation (Hakala, 2011). Hence, such orientations collectively can increase the effectiveness and efficiency of other orientations. Innovation could benefit from complementarity and the combination of SO leads to superior performance (e.g., Baker & Sinkula, 2009; Gnizy & Shoham, 2014; Ho et al., 2015; Mu & di Benedetto, 2011). Conversely, due to the overlapping of shared domains could dissipate their effects on innovation and firm performance (Baker & Sinkula, 2009).

Numerous studies have applied text analysis in organizational research. Particularly, in the context of strategic orientations, computer-aided text analysis (CATA) approach yielded significant contribution towards construct measurement and validation procedures of MO (Zachary, McKenny, Short, & Payne, 2011; Zachary, McKenny, Short, Davis, et al., 2011); EO (Engelen et al., 2015; McKenny, Short, et al., 2018; Short et al., 2009, 2010), and LO (Dutta et al., 2016). Dictionaries or word lists were developed and validated as well. Table 2 shows a comprehensive definition and operationalization of the variables involved and serves for a better understanding of the theoretical underpinning of the constructs involved.

METHOD

Text mining research provides the main framework for text analysis, as it encompasses the theoretical approaches, methods, techniques, and tools to promote the use of rich sourced data in the field (Kobayashi et al., 2018). Content analysis captures cognitions, emotions, and other types of meaning as reflected in the rhetoric presented in words or narrative texts (Short et al., 2018).

According to Indulska et al. (2012), within a text mining framework, it is possible to assume two approaches to conduct content analysis: conceptual and relational. In a conceptual analysis, text is examined for the presence of concepts; such concepts can represent words, phrases, or more complex constructs. Within this approach, the most representative application is computer-aided text analysis (CATA) which "enables the measurement of constructs by processing text into quantitative data based on the frequency of words" (McKenny, Aguinis, et al., 2018). In a conceptual analysis, algorithms read the text and classify concepts within the text into different categories based on dictionaries (Li, 2010).

Technically, a dictionary is a tabulated collection of items, each with an associated attribute, as, for example, in its traditional form of a word and associated definition. Thus, analysis is restricted to the term 'word lists,' where the created collections of words attempt to identify a particular attribute of a document (Loughran & Mcdonald, 2016). The occurrence of specific codes indicates the presence and salience of a construct of interest in the data (Pokorny et al., 2018; Short et al., 2018).

On the other hand, relational analysis approach pays attention not only to *what* is disclosed, but also to *how* it is disclosed. Co-occurrence is interpreted as an indicator of semantic proximity and refers to the above chance occurrence of two terms from a text corpus located in close proximity to each other in a certain order (Ignatow & Mihalcea, 2018). The fact that two concepts co-occur in an organizational document is interesting in itself: for instance, supposing that an annual report mentions innovation in every sentence in which customer is also mentioned. Even if the exact relation is unknown between the two concepts –if any–, it is known that the firm apparently associates customers with innovation, and this might influence readers if this association is strongly present in many annual reports.

Relationships between constituents of complex systems can be represented in terms of networks (Yang et al., 2016). Network analysis, from a graph theory approach, refers to the structure and visualization of individual entities regarded as nodes, and relationships or interactions between them, which are regarded as edges (Ignatow & Mihalcea, 2018). Nodes correspond to the constructs of interest into which text excerpts are coded and meaning is computationally explored, which in this paper refer to strategic orientations, innovation, and firm performance.

Groups of nodes highly connected between them but with few links to other nodes are called communities. These interconnected groups bring out much information about the network (Pons et al., 2005). Finding communities within a graph helps unveil the internal organization of a graph and can also be used to characterize the entities that compose it. Community detection extracts structural information of a network in an unsupervised manner, allowing to unveil the existence of a non-trivial internal network organization. This grouping method also let "to infer special relationships between the nodes that may not be easily accessible from direct empirical tests" (Yang et al., 2016).

Co-occurrence network analysis, the combination of the both previously mentioned approaches, provides a graphical visualization of the relationship between nodes –dimensions of SO, innovation and firm performance– extracted from texts –10-K annual reports–. Co-occurrence network analysis allows the discovery and visualization of the relationship patterns in the content of text collections (Matthies & Coners, 2015). Since concepts having similar appearance patterns are directly linked to one another, it is easier to identify the groups of concepts that represent main topics in texts and their centrality using a co-occurrence network, in comparison to other methods, such as multi-dimensional scaling and correspondence analysis (Higuchi, 2017).

Text mining procedure

KH Coder software for text mining was used to analyze the content of text sources and to explore the extracted information statistically and visually. It is a practical free software and its source code is open to the public (Higuchi, 2017). KH Coder is used in almost 500 scholarly publications (Deokar et al., 2018), and it is reviewed as a major text mining tool in marketing studies (Tang & Guo, 2015).

This paper conducts a co-occurrence network analysis following the steps for text mining proposed by Kobayashi et al. (2018) and Indulska et al. (2012), which refer on how to implement text analysis in an organizational research context.

The first step is related to the selection of text data, in this case annual reports, which are a source of rich and valuable data, from a qualitative point of view, which surveys or interviews cannot provide in the same manner.

The second step is text preprocessing, which includes 1) cleaning the text data retaining only the relevant text elements; 2) deleting unimportant characters –e.g., extra whitespaces, formatting tags–; 3) implementing a stop word removal procedure to ignore words which information content does not contribute to the meaning of the text, for instance, conjunctions and prepositions; and 4) parsing the text to obtain more efficient data, which implies extracting HTML code, embedded PDF's, and image items for creating compressed versions of the data. The purpose is to create a file in plain text which software can process to be transformed into mathematical structures –vectors and matrices–.

The third step refers to the text mining operations, which include conceptual and relational extraction through co-occurrence network analysis, for mining patterns of association of constructs. As result, a network map visualizes an undirected co-occurrence network where each construct represents a node in the network. The links or edges between nodes are represented by the magnitude of occurrences the two nodes have together.

Selection of text data: Annual reports on Form 10-K

Annual reports are considered the most important external document of any company as they contain crucial information about their financial performance and their future strategies (Kloptchenko et al., 2004). In the context of text analysis, annual reports are "prime materials to study the interaction of firms with their environment" (Duriau et al., 2007).

Methodologically, annual reports have several advantages in terms of reliability and exhaustiveness over other sources of corporate information to study cognitive phenomena (Duriau et al., 2007), and to obtain information on management's strategic posture. Moreover, annual reports are a valid and meaningful source of information about firm innovativeness and strategy (Michalisin, 2001). This has implications in a two-fold manner: 1) annual reports provide a common pool of knowledge to use relating to a company's strategy, products and services, risks, competitors, perspectives on the business results and what is driving them (U.S. Securities and Exchange Commision SEC, 2019); 2) this type of corporate narratives provides a valuable sampling frame for content analytic research because it maximizes sample size and increases the availability of texts from multiple time periods" (McKenny, Aguinis, et al., 2018).

As sampling frame, firms within materials and communications services from Standard & Poor's 500 companies (S&P 500) were selected. The S&P 500 lists the most valuable public companies in the U.S. and is widely regarded as the best single gauge of large-cap equities (S&P Dow Jones Indices, 2019). The index includes 500 leading companies and captures approximately 80% of the available market capitalization. Firms listed in S&P 500 are classified based on the Global Industry Classification Standard (GICS, 2019). In this case, two industry sectors –communication services and materials– are selected specifically because both represent exclusively services and manufacturing firms, respectively.

Communication services sector comprises "companies that facilitate communication and offer related content and information. It includes telecom, media and entertainment companies, producers of interactive gaming products and companies engaged in content and information creation or distribution through proprietary platforms" (MSCI Inc., 2018). World-wide known, young and most valued companies born in Silicon Valley such as Facebook, Netflix and Google belong to this industrial sector. On the other hand, the materials sector includes "companies that manufacture chemicals, construction materials, glass, paper, forest products, and related packaging products, and metals, minerals and mining companies, including producers of steel" (MSCI Inc., 2018).

The sample of texts comprises 140 selected firms' annual reports on Form 10-K from 2016 to 2018. The selected sample of firms consisted of 22 companies for the communication services sector and 26 for the materials sector, for a total of 48 companies (Table 1).

Table 1. Sample frame of S&P 500 companies within communication services and materials industry sectors (ordered by market-capitalization weightings)

NAME	SYMBOL	GICS SECTOR
Facebook, Inc.	FB	Communication Services
Alphabet Inc Class A	GOOGL	Communication Services
The Walt Disney Company	DIS	Communication Services
Verizon Communications	VZ	Communication Services
AT&T Inc.	T	Communication Services
Comcast Corp.	CMCSA	Communication Services
Netflix Inc.	NFLX	Communication Services
Linde PLC	LIN	Materials
DowDuPont	DWDP	Materials
Charter Communications	CHTR	Communication Services
Ecolab Inc.	ECL	Materials
Air Products & Chemicals Inc	APD	Materials

Activision Blizzard	ATVI	Communication Services
Sherwin-Williams	SHW	Materials
Electronic Arts	EA	Communication Services
Twitter, Inc.	TWTR	Communication Services
PPG Industries	PPG	Materials
LyondellBasell	LYB	Materials
Newmont Mining Corporation	NEM	Materials
Ball Corp	BLL	Materials
Weyerhaeuser	WY	Materials
International Paper	IP	Materials
CBS Corp.	CBS	Communication Services
Omnicom Group	OMC	Communication Services
Nucor Corp.	NUE	Materials
Freeport-McMoRan Inc.	FCX	Materials
Vulcan Materials	VMC	Materials
Twenty-First Century Fox Cl. A	FOXA	Communication Services
Intl Flavors & Fragrances	IFF	Materials
Celanese Corp.	CE	Materials
Martin Marietta Materials	MLM	Materials
Take-Two Interactive	TTWO	Communication Services
Eastman Chemical	EMN	Materials
CenturyLink Inc	CTL	Communication Services
FMC Corporation	FMC	Materials
Viacom Inc.	VIAB	Communication Services
CF Industries Holdings Inc	CF	Materials
WestRock	WRK	Materials
Avery Dennison Corp	AVY	Materials
Packaging Corporation A.	PKG	Materials
The Mosaic Company	MOS	Materials
Interpublic Group	IPG	Communication Services
Albemarle Corp	ALB	Materials
Dish Network	DISH	Communication Services
Sealed Air	SEE	Materials
TripAdvisor	TRIP	Communication Services
Discovery Inc. Class A	DISCA	Communication Services
News Corp. Class A	NWSA	Communication Services

Text preprocessing

Sample frame 10-K filings were obtained from The Notre Dame Software Repository for Accounting and Finance (SRAF, 2019). Originally, textual data is collected from the U.S. Securities and Exchange Commission (SEC, 2019) website. All the individual files –annual reports–were unified in a single plain text format in order to arrange a joint analysis. Stop word list from SRAF was used for the analysis.

Text mining operations

KH Coder allows content analysis using both a deductive conceptual –dictionary-based– and relational extraction approach. The software analyzes codes from the text data, using pre-defined dictionaries or word lists for the constructs of interest. Such dictionaries for coding SO, innovation and firm performance constructs were developed and validated by Dutta et al. (2016); McKenny, Short, et al. (2018); Short et al. (2010); Zachary, McKenny, Short, & Payne (2011). Paragraphs were the analysis unit. Table 2 shows the results of composing coding rules.

Table 2. Composing coding rule based on developed and validated dictionaries on market orientation, entrepreneurial orientation, learning orientation, innovation and firm performance

CONSTRUCT OF INTEREST	DIMENSION	CODING DEFINITION	EXAMPLE
Market orientation		The culture that effectively and efficiently creates value for customers (Narver & Slater, 1990) and the set of activities, processes and behaviors derived from the implementation of the marketing concept (Kohli and Jaworski, 1990).	
	Customer orientation	The adequate understanding of a customer's psyche as to provide "superior value" for said customer(s) in a continuous and sustainable manner (Narver & Slater, 1990)	In addition to creating new flavors and fragrances, our researchers and product development teams advise **customers** on ways to improve their existing products by adjusting or substituting current ingredients with more readily accessible or less expensive materials or by modifying the current ingredients to produce an enhanced yield. This often results in creating a better value proposition for our customers (International Flavors & Fragrances, 2016).

	Competitor orientation	The understanding of the short-term strengths and weaknesses as well as the long-term capabilities and strategies of both current and potential key competitors (Aaker, 1988; Day & Wensley, 1988; Porter, 1980, 1985)	**Competitors** often develop content that imitates or competes with our best-selling games and take sales away from them or reduce our ability to charge (Activision Blizzard, 2016).
	Interfunctional coordination	The coordination and utilization of a firm's resources, human or otherwise, to create "superior value" for the target buyer (Narver & Slater, 1990).	Our CRM model combines members of our team from within our manufacturing facilities and members of our business development team who reside remotely and nearer to our customers around the world. We also have **cross-functional teams** in the areas of quality, operational excellence, quoting, and design engineering with representatives from our various locations that provide support to our teams on a global basis (WestRock, 2018).
Entrepreneurial orientation		The specifically entrepreneurial aspects of firms' strategies to enact their organizational purpose, sustain its vision, and create competitive advantage involving the intentions, actions, processes, practices, and decision-making activities that lead to new entry and the pursuing of new market opportunities (Rauch et al., 2009; Lumpkin & Dess, 1996; Hakala, 2011; Covin & Slevin, 1989, 1991; Hult et al., 2004; Wiklund, 1999; Wiklund and Shepherd, 2005).	
	Autonomy	The actions of individuals or teams to surface and pursue opportunities to completion.	Our organization is highly **decentralized**, with most day-to-day operating decisions made by our division general managers and their staff (Nucor Corporation, 2018).
	Competitive aggressiveness	The aggressive organizational positioning or responses to defend against competitors, unfavorable industry trends, and other external threats.	We believe that we **compete** favorably on the factors described above. However, our industry is evolving rapidly and is becoming increasingly **competitive** (Twitter, 2016).
	Innovativeness	The willingness to encourage creativity and the development of new marketable ideas and inventions.	We endeavor to be the most **creative, innovative** and efficient company in our industry. Our core strategy is to capitalize on the popularity of video games by developing and publishing high-quality interactive entertainment experiences across a range of genres (Take-Two Interactive, 2018).
	Proactiveness	The anticipation of future changes and the undertaking of appropriate, often innovative, action to capitalize on the opportunity or mitigate the threat.	Our failure to effectively **anticipate** or adapt to new technologies and changes in consumer expectations and behavior could significantly adversely affect our competitive position and our business and results of operations (Charter Communications, 2017).
	Risk-taking	The willingness to take bold action in the face of uncertainty.	We **face risks** relating to competition for the leisure time and discretionary spending of audiences, which has intensified in part due to advances in technology and changes in consumer expectations and behavior (Charter Communications, 2018).
Learning orientation		The key values that influences the propensity of the firm to learn by generating, processing and using market information and new knowledge in order to gain competitive advantage (Sinkula et al., 1997; Calantone et al., 2002).	
	Commitment to learning	The organizational value toward learning, which influences the intensity to promote a learning culture (Sinkula, Baker, & Noordewier, 1997).	We invest substantial capital in our content, including in the production of original content on our networks, in our films and in our television production business, before **learning** the extent to which it will garner critical success and popularity with consumers (Viacom, 2018).
	Open-mindedness	The willingness to critically evaluate the operational routine and accept new ideas (Sinkula, et al., 1997).	We take great pride in our culture. We embrace collaboration and creativity and encourage the iteration of **ideas** to address complex technical challenges. Transparency and **open dialogue** are central to how we work, and we like to ensure that company news reaches our employees first through internal channels (Google, 2018).
	Shared vision	The focus or direction of learning among the members of an organization. Without a shared vision, individuals are less likely to know what organizational expectations exist, what outcomes to measure, or what theories in use are in operation (Sinkula, et al., 1997).	All significant events are investigated, and lessons learned are **shared** with workers (Newmont Mining Corporation, 2017).
Innovation (as an outcome)		The consequences of innovation activities or the outputs of innovation process (Crossan & Apaydin, 2010).	The timely **introduction of new products** and **improvements** in current products helps determine our success (Avery Dennison Corp., 2016).
Firm performance		The economic outcomes resulting from the interplay among an organization's attributes, actions, and environment (Combs, Crook, and Shook 2005, p. 262) capturing the underlying manifestations of how well a firm is effectively satisfying its stated goals (Bergh et al., 2016; Combs et al., 2005).	Changes in our business strategy or restructuring of our businesses may increase our costs or otherwise affect the profitability of our businesses (Walt Disney Company, 2017).

To identify the presence and relevance of constructs of interest, KH Coder conducts a frequency analysis of terms included in the coding rule –dictionaries–. It results in a frequency list indicating the number of paragraphs each code applies to, and its percentage of the total.

To extract patterns of association of constructs, a co-occurrence network analysis was conducted. KH Coder identifies the relationships between constructs using the Jaccard similarity coefficient, which is a parameter used to compare characteristic similarity and proximity between sets of information efficiently without the use of data redundancy (Irani et al., 2016; Singthongchai & Niwattanakul, 2013). Centrality is reflected by the influence of a construct in texts and determines what kind of role it plays in a textual network. The degree of centrality is manifested in terms of the number of nodes to which a given node is directly connected (Higuchi, 2017).

KH Coder detects communities structure using the Walktrap algorithm (Pons et al., 2005). Additionally, analysis of the minimum spanning tree (MST) was provided, based on the strength of co-occurrence using the Prim method (Higuchi, 2017). MST indicates which associations are most important in the network. Mathematically, is defined as the sub-network that connects all nodes while minimizing the link weights and without forming loops (Tewarie et al., 2015).

RESULTS

Co-occurrence network maps –Figures 1 and 2– are the outputs derived from text mining analysis, visualizing the patterns of association. Table 3 lists the results of the obtained frequency list, which quantitatively shows the relative relevance of composite strategic orientations, innovation and firm performance reflected in the annual reports. Table 4 lists strategic orientations in a multidimensional view, showing the relative relevance of each individual dimension in the text. Jointly, almost half of the frequencies (42.65%) belong to MO (19.76%), EO (15.54%) and LO (6.97%), which indicate that narrative exhibitions on these constructs are significantly important for S&P companies in terms of their business and strategy. A total of 102,051 paragraphs were analyzed.

Figure 1 and 2 shows the aggregated and disaggregated view of strategic orientations, respectively, and their relationships with innovation and firm performance. The size of nodes represents the relative frequency of constructs –as shown in Table 2– and the Jaccard distances –coefficients of the edges– indicate the relative degree of their co-occurrence, that is, the strength of connections between them. The network map is represented through minimum span tree, in which all nodes are connected to each other directly or indirectly to indicate substantive relationships among constructs of interest.

Table 3. Frequency list of constructs: strategic orientations as composite constructs, innovation, and firm performance

CONSTRUCTS OF INTEREST	FREQUENCY	PERCENT
Market_Orientation	20’165	19.76%
Entrepreneurial_Orientation	15’860	15.54%
Learning_Orientation	7108	6.97%
Innovation	2286	2.24%
Firm_Performance	20’795	20.38%
N of Paragraphs	102’051	

Table 4. Frequency list of constructs: strategic orientations' dimensions, innovation, and firm performance

CONSTRUCTS OF INTEREST	FREQUENCY	PERCENT
MO_CompetitorOrient	6180	6.06%
MO_CustomerOrient	11’912	11.67%
MO_InterfuncCoord	6884	6.75%
EO_Autonomy	1211	1.19%
EO_CompetitiveAggressiveness	5599	5.49%
EO_Innovativeness	6111	5.99%
EO_Proactiveness	5527	5.42%
EO_RiskTaking	1241	1.22%
LO_CommitmentLearning	4563	4.47%
LO_OpenMindedness	720	0.71%
LO_SharedVision	2103	2.06%
N of Paragraphs	102’051	

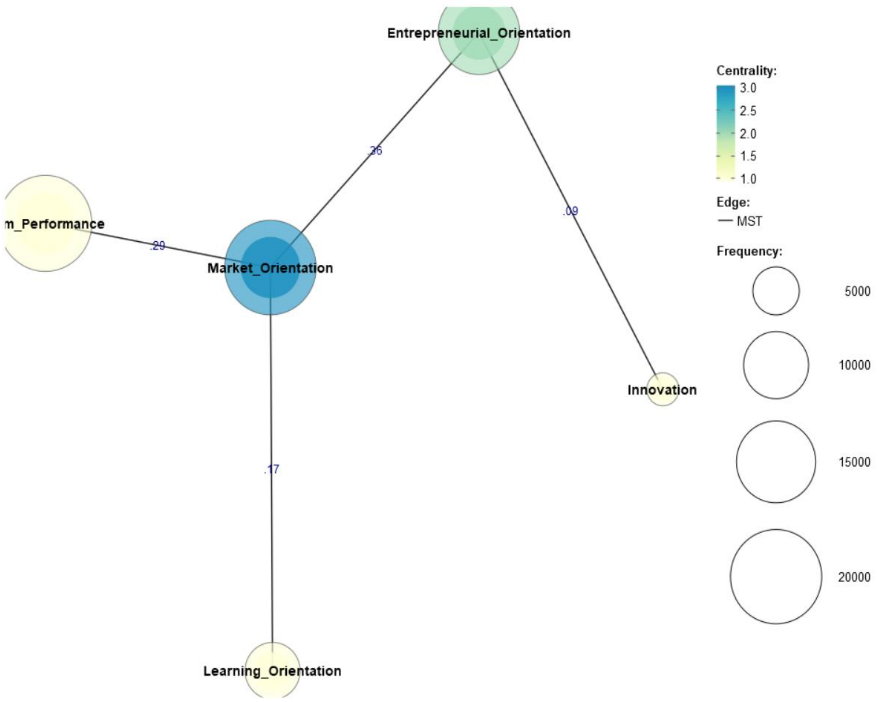

Figure 1. Co-occurrence network map of composite strategic orientations, innovation and firm performance

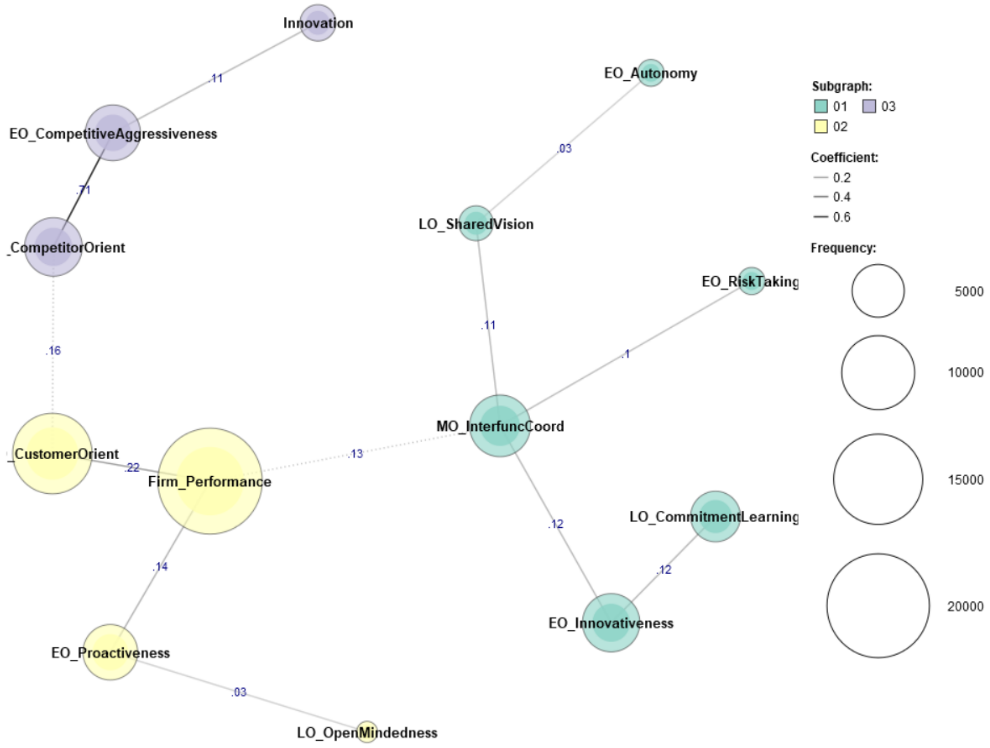

Figure 2. Co-occurrence network map of strategic orientations' dimensions, innovation and firm performance

As shown in Figure 1, the aggregated view of strategic orientations and the links with innovation and firm performance, MO plays a central role in corporate narratives. This is evidenced by its multiple direct links with EO, LO and Firm Performance which is the most exhibited construct in narratives. Also, EO is linked to innovation. In this sense, Innovation is indirectly linked to Firm Performance through EO and MO. LO is also indirectly linked to Firm Performance through MO.

As shown in Figure 2, the Walktrap community detection algorithm generates three visually colored communities as follows:

Community 1 includes Firm Performance, MO-Customer Orientation, EO-Proactiveness, and LO-Open Mindedness. Narratives on Firm performance are directly and more associated to MO-Customer Orientation and EO-Proactiveness. EO-Proactiveness and LO-Open Mindedness are less associated.

Community 2 includes MO–Competitor Orientation, EO-Competitive Aggressiveness, and Innovation. Narratives on EO-Competitive Aggressiveness are directly and strongly associated with MO–Competitor Orientation, and Innovation.

Community 3 includes MO-Interfunctional Coordination, EO-Innovativeness, LO-Shared Vision, LO-Commitment to Learning, EO-Autonomy, and EO-Risk Taking. MO-Interfunctional Coordination is more associated to EO-Innovativeness, LO-Shared Vision, and EO-Risk Taking. EO-Innovativeness is also associated to LO-Shared Vision. LO-Shared Vision is associated to EO-Autonomy.

Communities are interconnected. Communities 1 and 2 are linked by MO-Customer Orientation and MO–Competitor Orientation. Communities 1 and 3 are linked by Firm Performance and MO-Interfunctional Coordination.

DISCUSSION AND IMPLICATIONS

This exploratory data-driven paper applied co-occurrence network analysis on a sample of S&P 500 companies' 10-K annual reports in order to explore the exhibited relationships between strategic orientations, innovation and firm performance and to extract their co-occurrence patterns, for a better understanding how these constructs of interest are adopted, combined and balanced in existing business practices through text mining rich and publicly accessible data sources.

By assuming both an aggregated –composite of discrete but related set of dimensions– and disaggregated –individual dimensions– approaches of the strategic orientation this paper facilitated a simpler explanations from more parsimonious view of the relationships on overall text data; and on the other, more complex relationships, avoiding 'excessive aggregation' and 'aiding prediction' (McKenny, Short, et al., 2018).

From an aggregated view of strategic orientations, it is demonstrated the relevance and centrality of these constructs of interest among corporate disclosures. Results indicated that MO plays a central role in the relationships between SO, innovation and firm performance, supporting the idea that MO has become a cost of doing business, in order to prevent business failure (Kumar et al., 2011).

Although firm performance's narratives are the most exhibited in annual reports –as expected since corporate disclosures aims to provide overview information on business and financial condition–, MO is the construct which is more connected with alternative orientations, innovation and firm performance. These findings support past research in the sense that firms are more likely to associate MO with LO or EO (Grinstein, 2008). EO and LO play a supporting role in creating value for customers by pursuing the right market opportunities and influencing the creation and use of knowledge and insights needed to capitalize on these opportunities.

Surprisingly, narratives related to innovation outcomes, such as the introduction of new products or services into markets, are not closely related to the businesses results' narratives. The former is more associated to creativity and exploration of market opportunities rather than the latter one on specific financial and market results.

From a SO disaggregated view, this study found three major co-occurrence patterns of association between strategic orientations, innovation and firm performance that represent firms' narrative exhibitions with regards to their business and strategy. Overall, customers, competitors, and resources within firms cover a central place on organizational narratives exhibitions.

First, companies place great emphasis on associating aspects such as firm's profitability, finance, sales, reputation, and other goals results with the adequate understanding of current and future customers' needs to provide them with superior value, and with the anticipation and capitalization of market opportunities. Narratives on anticipating future changes or mitigating threats are slightly connected to accepting new ideas and questioning operative routines.

Second, competition has a special emphasis in corporate narratives. Companies associate the understanding of weaknesses and strengths of competitors and the responses to defend against them, industry trends and external threats with leveraging the introduction of new products/services into markets. It comes into sight that first-mover advantage leads to a better defense against increasing competition.

Third, companies further exhibit the importance of synergies developed by the different functional areas working together to improve creativity and innovation processes and with shared learning expectations among individuals and teams. Creative and explorative firms' exhibitions are also associated with the promotion of a learning culture. The coordination and utilization of resources are associated with encouraging employees to take bold actions to venture into uncertain outcomes. Narratives on organizational expectations about learning are slightly connected with individual and teams' autonomous actions to pursue opportunities.

Although the three SO reviewed are different constructs, they can act complementarily and simultaneously. Practitioners may find useful to acknowledge that fostering and exhibiting market, entrepreneurial and learning orientations within their firms could lead to enhance innovations outcomes and achieve superior performance. In this sense, putting customer satisfaction at the center of the firm's activity, improving the quality of learning from external environment and pursuing new market opportunities through the development of new products or services should lead to gain competitive advantage and enjoy superior firm performance superior performance in comparison with competitors. The synergies created by the adoption and combination of the aforementioned activities are exhibited by highly successful companies, such as listed in S&P 500, as demonstrated in this study.

Researchers may find useful to mine qualitative rich and in-depth public text data to unveil underlying organizational phenomena of interest provided by public institutions, customers, markets and other interesting sources. That could not be possible to analyze with other common research methods (Kobayashi et al., 2018). Although for more exploratory or descriptive studies applying text mining is not mandatory to establish the validity of inferences (Short et al., 2009), the conclusions of this study must be interpreted in light of their limitations. For example, while text mining procedures can identify words and phrases associated with constructs of interest, it cannot interpret the use of this language in context, which can lead to misinterpretations (McKenny, Short, et al., 2018). Still, co-occurrence of constructs is a strong indication of the presence, relevance and resilience of constructs of interest in organizational narratives.

Future research, from a contingency approach, could focus on differences between various types of firms in order to analyze whether the patterns of combinations persist or not regardless of contextual moderators such as firm size, industry sector and national culture. Subgroup analysis could provide a better understanding of phenomena under study.

REFERENCES

Baker, W. E., & Sinkula, J. M. (2009). The complementary effects of market orientation and entrepreneurial orientation on profitability in small businesses. *Journal of Small Business Management*, *47*(4), 443–464. https://doi.org/10.1111/j.1540-627X.2009.00278.x

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Bergh, D. D., Aguinis, H., Heavey, C., Ketchen, D. J., Boyd, B. K., Su, P., Lau, C. L. L., & Joo, H. (2016). Using meta-analytic structural equation modeling to advance strategic management research: Guidelines and an empirical illustration via the strategic leadership-performance relationship. *Strategic Management Journal*, *37*(3), 477–497. https://doi.org/10.1002/smj.2338

Calantone, R. J., Cavusgil, S. T., & Zhao, Y. (2002). Learning orientation, firm innovation capability, and firm performance. *Industrial Marketing Management*, *31*(6), 515–524. https://doi.org/10.1016/S0019-8501(01)00203-6

Cambra-Fierro, J. J., Hart, S., Polo-Redondo, Y., & Fuster-Mur, A. (2012). Market and learning orientation in times of turbulence: relevance questioned? An analysis using a multi-case study. *Quality & Quantity*, *46*(3), 855–871. https://doi.org/10.1007/s11135-011-9429-x

Castelfranchi, Y. (2017). Computer-aided text analysis: an open-aired laboratory for social sciences. *Journal of Science Communication*, *16*(02), 1–11. https://doi.org/10.22323/2.16020304

Combs, J. G., Crook, T. R., & Shook, C. L. (2005). The Dimensionality of Organizational Performance and its Implications for Strategic Management Research. In *Research Methodology in Strategy and Management -Volume 2* (Vol. 2, pp. 259–286). https://doi.org/10.1016/S1479-8387(05)02011-4

Covin, J. G., & Slevin, D. P. (1991). A Conceptual Model of Entrepreneurship as Firm Behavior. *Entrepreneurship Theory and Practice*, *16*(1), 7–26. https://doi.org/10.1177/104225879101600102

Crossan, M. M., & Apaydin, M. (2010). A multi-dimensional framework of organizational innovation: A systematic review of the literature. *Journal of Management Studies*, *47*(6), 1154–1191. https://doi.org/10.1111/j.1467-6486.2009.00880.x

Delen, D., & Crossland, M. D. (2008). Seeding the survey and analysis of research literature with text mining. *Expert Systems with Applications*, *34*(3), 1707–1720. https://doi.org/10.1016/j.eswa.2007.01.035

Deokar, A. v, Gupta, A., Iyer, L. S., & Jones, M. C. (2018). *Analytics and Data Science* (A. v. Deokar, A. Gupta, L. S. Iyer, & M. C. Jones, Eds.). Springer International Publishing. https://doi.org/10.1007/978-3-319-58097-5

Deshpande, R., & Webster, F. E. (1989). Organizational Culture and Marketing: Defining the Research Agenda. *Journal of Marketing*, *53*(1), 3–15. https://doi.org/10.2307/1251521

Deutscher, F., Zapkau, F. B., Schwens, C., Baum, M., & Kabst, R. (2016). Strategic orientations and performance: A configurational perspective. *Journal of Business Research*, *69*(2), 849–861. https://doi.org/10.1016/j.jbusres.2015.07.005

Duriau, V. J., Reger, R. K., & Pfarrer, M. D. (2007). A content analysis of the content analysis literature in organization studies: Research themes, data sources, and methodological refinements. *Organizational Research Methods*, *10*(1), 5–34. https://doi.org/10.1177/1094428106289252

Dutta, D. K., Gupta, V. K., & Chen, X. (2016). A Tale of Three Strategic Orientations: A Moderated-Mediation Framework of the Impact of Entrepreneurial Orientation, Market Orientation, and Learning Orientation on Firm Performance. *Journal of Enterprising Culture*, *24*(3), 313–348. https://doi.org/10.1142/S0218495816500126

Engelen, A., Neumann, C., & Schwens, C. (2015). "Of Course I Can": The Effect of CEO Overconfidence on Entrepreneurially Oriented Firms. *Entrepreneurship: Theory and Practice*, *39*(5), 1137–1160. https://doi.org/10.1111/etap.12099

Filippov, S., & Hofheinz, P. (2016). Text and Data Mining for Research and Innovation. In *The Lisbon Council* (Issue 20).

Gatignon, H., Gotteland, D., & Haon, C. (2016). Strategic and Market Orientations. In *Making Innovation Last: Volume 1* (pp. 97–152). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-137-56098-8_4

Gatignon, H., & Xuereb, J.-M. (1997). Strategic orientation of the firm and new product performance. *Journal of Marketing Research*, *34*(1), 77-90. 14p. 2 Diagrams. https://doi.org/10.2307/3152066

Gnizy, I., Baker, W., & Grinstein, A. (2014). Proactive learning culturee: A dynamic capability and key success factor for SMEs entering foreign markets. *International Marketing Review*, *31*(5), 477–505. https://doi.org/10.1108/IMR-10-2013-0246

Gnizy, I., & Shoham, A. (2014). Uncovering the influence of the international marketing function in international firms. *International Marketing Review*, *31*(1), 51–78. https://doi.org/10.1108/IMR-12-2012-0204

Grinstein, A. (2008). The relationships between market orientation and alternative strategic orientations: A meta-analysis. *European Journal of Marketing*, *42*(1/2), 115–134. https://doi.org/10.1108/03090560810840934

Hakala, H. (2011). Strategic Orientations in Management Literature: Three Approaches to Understanding the Interaction between Market, Technology, Entrepreneurial and Learning Orientations. *International Journal of Management Reviews*, *13*(2), 199–217. https://doi.org/10.1111/j.1468-2370.2010.00292.x

Higuchi, K. (2017). A Two-Step Approach to Quantitative Content Analysis: KH Coder Tutorial using Anne of Green Gables (Part I). *Ritsumeikan Social Sciences Review*.

Ho, J., Plewa, C., & Lu, V. N. (2015). Examining strategic orientation complementarity using multiple regression analysis and fuzzy set QCA. *Journal of Business Research*, *69*(6), 2199–2205. https://doi.org/10.1016/j.jbusres.2015.12.030

Hult, G. T. M., Hurley, R. F., & Knight, G. A. (2004). Innovativeness: Its antecedents and impact on business performance. *Industrial Marketing Management*, *33*(5), 429–438. https://doi.org/10.1016/j.indmarman.2003.08.015

Hult, G. T. M., & Ketchen, D. J. (2001). Does market orientation matter?: A test of the relationship between positional advantage and performance. *Strategic Management Journal*, *22*(9), 899–906. https://doi.org/10.1002/smj.197

Ignatow, G., & Mihalcea, R. (2018). *An introduction to text mining: research design, data collection, and analysis*.

Indulska, M., Hovorka, D. S., & Recker, J. (2012). Quantitative approaches to content analysis: Identifying conceptual drift across publication outlets. *European Journal of Information Systems*, *21*(1), 49–69. https://doi.org/10.1057/ejis.2011.37

Irani, J., Pise, N., & Phatak, M. (2016). Clustering Techniques and the Similarity Measures used in Clustering: A Survey. *International Journal of Computer Applications*, *134*(7), 9–14. https://doi.org/10.5120/ijca2016907841

Kloptchenko, A., Eklund, T., Karlsson, J., Back, B., Vanharanta, H., & Visa, A. (2004). Combining data and text mining techniques for analysing financial reports. *Intelligent Systems in Accounting, Finance & Management*, *12*(1), 29–41. https://doi.org/10.1002/isaf.239

Kobayashi, V. B., Mol, S. T., Berkers, H. A., Kismihók, G., & den Hartog, D. N. (2018). Text Mining in Organizational Research. *Organizational Research Methods*, *21*(3), 733–765. https://doi.org/10.1177/1094428117722619

Kozlenkova, I. v., Samaha, S. A., & Palmatier, R. W. (2014). Resource-based theory in marketing. *Journal of the Academy of Marketing Science*, *42*(1), 1–21. https://doi.org/10.1007/s11747-013-0336-7

Leenders, M., Farrell, M. A., & van der Wurff, R. (2016). Market or society? Dual orientations and the impact on innovativeness in media organizations. *Journal of Strategic Marketing*, *4488*(April), 1–15. https://doi.org/10.1080/0965254X.2016.1149209

Li, F. (2010). Textual Analysis of Corporate Disclosures: A Survey of the Literature. *Journal of Accounting Literature*, *29*, 143.

Lonial, S. C., & Carter, R. E. (2015). The Impact of Organizational Orientations on Medium and Small Firm Performance: A Resource-Based Perspective. *Journal of Small Business Management*, *53*(1), 94–113. https://doi.org/10.1111/jsbm.12054

Loughran, T., & Mcdonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, *54*(4), 1187–1230. https://doi.org/10.1111/1475-679X.12123

Lumpkin, G. T., & Dess, G. G. (1996). Clarifying the Entrepreneurial Orientation Construct and Linking It to Performance. *Academy of Management Review*, *21*(1), 135–172. https://doi.org/10.2307/258632

Matthies, B., & Coners, A. (2015). Computer-Aided Text Analysis of Corporate Disclosures - Demonstration and Evaluation of Two Approaches. *The International Journal of Digital Accounting Research*, *15*, 69–98. https://doi.org/10.4192/1577-8517-v15

McKenny, A., Aguinis, H., Short, J., & Anglin, A. (2018). What Doesn't Get Measured Does Exist: Improving the Accuracy of Computer-Aided Text Analysis. *Journal of Management*, *44*(7), 2909–2933. https://doi.org/10.1177/0149206316657594

McKenny, A., Short, J., Ketchen, D., Payne, G., & Moss, T. (2018). Strategic entrepreneurial orientation: Configurations, performance, and the effects of industry and time. *Strategic Entrepreneurship Journal*, *12*(4), 504–521. https://doi.org/10.1002/sej.1291

Michalisin, M. D. (2001). Validity of annual report assertions about innovativeness: An empirical investigation. *Journal of Business Research*, *53*(3), 151–161. https://doi.org/10.1016/S0148-2963(99)00118-6

MSCI Inc. (2018). *Global Industry Classification Standard Methodology* (Issue September). https://www.msci.com/documents/1296102/11185224/GICS+Sector+definitions+Sept+2018.pdf/afc87e7b-bbfe-c492-82af-69400ee19e4f

Mu, J., & di Benedetto, C. A. (2011). Strategic orientations and new product commercialization: Mediator, moderator, and interplay. *R & D Management*, *41*(4), 337–359. https://doi.org/10.1111/j.1467-9310.2011.00650.x

Pokorny, J. J., Norman, A., Zanesco, A. P., Bauer-Wu, S., Sahdra, B. K., & Saron, C. D. (2018). Network analysis for the visualization and analysis of qualitative data. *Psychological Methods*, *23*(1), 169–183. https://doi.org/10.1037/met0000129

Pollach, I. (2012). Taming textual data: The contribution of corpus linguistics to computer-aided text analysis. *Organizational Research Methods*, *15*(2), 263–287. https://doi.org/10.1177/1094428111417451

Pons, P., Latapy, M., & Pons, P. (2005). Computing communities in large networks using random walks (long version). In *Computer and Information Sciences - ISCIS 2005*. https://doi.org/10.1007/11569596

Short, J., Broberg, J., Cogliser, C., & Brigham, K. (2010). Construct Validation Using Computer-Aided Text Analysis (CATA): An Illustration Using Entrepreneurial Orientation. *Organizational Research Methods*, *13*(2), 320–347. https://doi.org/10.1177/1094428109335949

Short, J., McKenny, A., & Reid, S. (2018). More Than Words? Computer-Aided Text Analysis in Organizational Behavior and Psychology Research. *Annual Review of Organizational Psychology and Organizational Behavior*, *5*(1), 415–435. https://doi.org/10.1146/annurev-orgpsych-032117-104622

Short, J., Payne, G., Brigham, K., Lumpkin, G., & Broberg, J. (2009). Family Firms and Entrepreneurial Orientation in Publicly Traded Firms: A Comparative Analysis of the S&P 500. *Family Business Review*, *22*(1), 9–24. https://doi.org/10.1177/0894486508327823

Singthongchai, J., & Niwattanakul, S. (2013). A Method for Measuring Keywords Similarity by Applying Jaccard's, N-Gram and Vector Space. *Lecture Notes on Information Theory*, *1*(4), 159–164. https://doi.org/10.12720/lnit.1.4.159-164

Sinkula, J. M., Baker, W. E., & Noordewier, T. (1997). A Framework for Market-Based Organizational Learning: Linking Values, Knowledge, and Behavior. *Journal of the Academy of Marketing Science*, *25*(4), 305–318. https://doi.org/10.1177/0092070397254003

S&P Dow Jones Indices. (2019). *S&P Dow Jones Indices: Index Methodology* (Issue April). https://us.spindices.com/indices/equity/sp-500

Tang, C., & Guo, L. (2015). Digging for gold with a simple tool: Validating text mining in studying electronic word-of-mouth (eWOM) communication. *Marketing Letters*, *26*(1), 67–80. https://doi.org/10.1007/s11002-013-9268-8

Tewarie, P., van Dellen, E., Hillebrand, A., & Stam, C. J. (2015). The minimum spanning tree: An unbiased method for brain network analysis. *NeuroImage*, *104*, 177–188. https://doi.org/10.1016/j.neuroimage.2014.10.015

The Notre Dame Software Repository for Accounting and Finance (SRAF). (2019). *Software Repository for Accounting and Finance*. https://sraf.nd.edu/

U.S. Securities and Exchange Commision. (2019). *SEC.gov | Form 10-K*. Fast Answers - Form 10-K. https://www.sec.gov/fast-answers/answers-form10khtm.html

Yang, Z., Algesheimer, R., & Tessone, C. J. (2016). A Comparative Analysis of Community Detection Algorithms on Artificial Networks. *Scientific Reports*, *6*(1–16). https://doi.org/10.1038/srep30750

Zachary, M., McKenny, A., Short, J., Davis, K., & Wu, D. (2011). Franchise branding: An organizational identity perspective. *Journal of the Academy of Marketing Science*, *39*(4), 629–645. https://doi.org/10.1007/s11747-011-0252-7

Zachary, M., McKenny, A., Short, J., & Payne, G. (2011). Family business and market orientation: Construct validation and comparative analysis. *Family Business Review*, *24*(3), 233–251. https://doi.org/10.1177/0894486510396871

Zhou, K. Z., Yim, C. K., & Tse, D. K. (2005). The Effects of Strategic Orientations on Technology- and Market-Based Breakthrough Innovations. *Journal of Marketing*, *69*(2), 42–60. https://doi.org/10.1509/jmkg.69.2.42.60756