HUMAN CAPITAL OF UPPER ECHELONS: THEIR INFLUENCE ON STRATEGIC CHANGE DECISIONS INVOLVING DIGITAL TRANSFORMATION

By:

TUSHAR RAVINDRA SHAH

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Management at The University of Texas at Arlington August, 2020

Arlington, Texas

Supervising Committee:

Dr. George Benson, Chair
Dr. Sridhar Nerur, Member
Dr. Jeffrey McGee, Member

ABSTRACT

HUMAN CAPITAL OF UPPER ECHELONS: THEIR INFLUENCE ON STRATEGIC CHANGE DECISIONS INVOLVING DIGITAL TRANSFORMATION

TUSHAR RAVINDRA SHAH

Supervising Professor: George S. Benson

My dissertation studies strategic change in the context of digital transformation and the role of organization's upper echelons (specifically the human capital of TMTs and CEOs) in predicting digital transformation. I concur with the prevalent belief that digital transformation is the fourth industrial revolution. It is already disrupting several businesses and business models as a result of which existing models of strategic change may have to be modified to explain these drastic transformations. I begin with defining the domain of strategic change as experienced by organizations when they intend to or are undergoing digital transformation. To support my theoretical assertions, I use upper echelons theory, multi-lens model of strategic change and dynamic capabilities perspective, in combination with human capital literature. Using text analytics, I empirically derive a measure for digital transformation. I then test some of my hypotheses related to different human capital characteristics of top management team members and CEOs of the top 500 organizations as per S&P 500 rankings. I also examine the moderating influence of environmental dynamism in the relationship between upper echelons characteristics and strategic change related to digital transformation. I run two different models using multilevel (hierarchical) linear regression and present a composite picture of my findings. I find that TMT role heterogeneity is positively related to digital transformation. This relationship is moderated

such that at high dynamism, it remains positive whereas for low dynamism it becomes non-significant. I further find that digital transformation follows an inverse U-shaped curvilinear relationship with TMT mean tenure such that it is lowest at low and high levels of TMT mean tenure. In moderating conditions, this relationship's magnitude increases under low dynamism but becomes non-significant under high dynamism. CEO duality is negatively associated whereas CEO share-owning is positively associated with digital transformation, and finally CEO tenure is negatively associated with digital transformation under low dynamism but non-significant at high dynamism.

My main proposed contribution to literature is to supplement the existing models of strategic change in the context of technological changes in organizations. Secondly, I use some unique methods and techniques to define the term digital transformation.

Key Words: Digital Transformation, Upper Echelons, TMT, CEO, Text Analytics, Strategic Change

ACKNOWLEDGEMENTS

Aum Shri Ganeshaaya Namah! Aum Shri Saraswatyaye Namah!! Aum Shri Guravay Namah!!!

Any literary work or any auspicious work related to Sciences that anyone embarks upon in Indian Culture is commenced only after invocation to Lord Ganesha - The Presiding Deity of Inaugurations, The Lord of Intellect and Wisdom, The Patron of Arts and Sciences, and finally the Destroyer of Obstacles. My prayers to Him for success in all my endeavors. The intellect and wisdom inspired by Lord Ganesha can be fruitfully applied with the endowments facilitated by Goddess Mother Saraswati, the Goddess of Learning. She inspires acquisition and processing of new knowledge on an ongoing basis. But both these deities are not present in their gross form. They manifest themselves through the Guru – the remover of darkness of ignorance.

Guroor Brahma, Guroor Vishnu, Guroor Devo Maheshwara, Guruh Saakshaat Parabrahmah, Tasmyee Shri Guravay Namah!

The Guru is the Brahma – The Creator; Guru is the Vishnu – The Preserver; Guru is the Shiva or Maheshwara – The Destroyer. The Guru is the embodiment of what some people refer to as God – i.e., The Unitary Ultimate Principle, The Divine Consciousness, and The Cosmic Energy. Thus, I pray to the Guru to seek His Perpetual Blessings!

And so, all the first acknowledgements go to different embodiments of the Guru – all the professors and teachers.

Dr. Benson, my Chair, my friend, and guide. My deepest gratitude to him for his support of my candidature to the PhD program. My humblest prostrations to him for being there with me

throughout the program, handholding me through the labyrinth of the academic world. His ever-cheerful demeanor has always raised my spirits even without being drunk. His infectious enthusiasm is also very contagious, something positive about contagion in the current times! Despite being head-deep in responsibilities of the department, he has always found time for me to spend on the problems that I took to him. He has always been a sounding board for me during my personal, academic, and professional challenges.

I am very thankful to Dr. Nerur, committee member, for being so active in helping me out with text analytics related work in all the previous projects and with my dissertation. Despite being neck-deep in so many multiple activities, he has always provided me with solutions to my queries related to my areas of interest in which he is an expert.

I thank Dr. McGee for agreeing to be a part of my committee despite not having known me well prior to that. I am also grateful to him for being supportive of my candidature at my prospective employment.

I am grateful to Dr. Datta for being a part of my committee right until his retirement. My present dissertation has been crafted based on some of the important inputs that he provided during my interaction with him. My prayers and best wishes to him for a great retired life.

Dr. Casper has been a source of inspiration for all of us. Her nurturing affection for all of us PhD students has always encouraged us to persist strongly in the program despite several challenges – political environment-related immigration issues, unfortunate mishaps during our tenure, and career and job-related hurdles. Throughout our stay, her solid support for us has been relentless. She has gone out of her way to make us personally, professionally, and academically as comfortable as possible.

My thanks to Dr. Bell for her encouragement and inspiration whenever we interacted with her. As a teacher and as an advisor, she has been very supportive to all that we have intended to do during our stay.

All the teachers who taught me – Dr. McFadyen, Dr. Grisaffe, Dr. Lavelle, Dr. Yasar; have all been instrumental in allaying my darkness of ignorance and enlightened me towards the path I intended to tread on. I am grateful to all of them for sharing their knowledge with me and my cohorts.

Amongst all my professors, the deepest gratitude is reserved for one person – the one who has had a lasting impression on me and the career path I wish to take henceforth. No amount of space can justify my acknowledgement and dedication to him – one and only Dr. Rasheed. In the year 2015 when I just landed in Arlington directly from India, I was grappling with several challenges. I could no longer work as I was on a dependent status; I was managing my family as a stay-at-home dad when my wife used to go for work. That is when I met him. Despite being my very first time and despite me being a total stranger to him, Dr. Rasheed who was also the Chair of the department at that time, spent one full hour with me, helping me understand what I could do in my situation. It is then that I decided that I have to pursue my PhD nowhere else but in the Management Department of the University of Texas at Arlington. After I started in the program, his wise counsel has always been the most sought-after for all of us. Dr. Rasheed has spent hours and hours guiding me and counseling me on professional, academic, and more importantly personal fronts. Last 4 years have been nothing short of a roller-coaster ride for me at a personal level. My family and I have faced some of the most difficult times of our life and every time Dr. Rasheed has been the pillar of strength for me. After spending time with him on every difficult situation, I have always walked out of his room with most encouraging thoughts

and many a times, solutions too. Academically and professionally, his wisdom has always prevailed upon all the challenges I faced and decisions I have taken. Once again, my most respectful salutations to Dr. Rasheed!

My next set of acknowledgements are reserved for my co-students and cohorts. Ajith, Ankita, and Joohan have been very supportive during all these years. We have spent together the most exhilarating as well as frustrating times, and every time we have been enjoying together or helping each other out. Special thanks to Ajith for guiding us through the various unknown academic variables. He has helped us with understanding literature, he has helped us with understanding methods. He has helped us think beyond.

Thanks are due for all our seniors – Ryan, Hoda, Anna, Kilho, Faezeh, Nitin, and Manisha. All of them have helped us shape up to the challenges of the program with their own personal examples, experience, and academic guidance. My special note of thanks are reserved for my brother from another mother – Marwan. He has always gone much beyond with all the guidance he has provided to me. Helping me understand concepts, involving me in projects, guiding me in my career aspirations, and loving me unconditionally. His insights into several subjects of my interest have helped me shape my paradigms about the fields of management and strategy.

I thank all my immediate juniors, Esther, Ifey, Mike, and Shelia for sharing their thoughts, views and ideas during the seminar courses that we took together. This has enriched my knowledge and broadened my horizons. I also thank my super-juniors – Dan and Marla. All my interactions with them have been value-adding for me, personally as well as academically.

The whole of Management Department, my “karmabhoomi”, my abode of karma or action deserves all the credit for hosting me all through these years – for providing me with all the support I need, for providing me all the infrastructure I wanted. The professors, the office staff and the student assistants have all been very helpful in facilitating my progress through the years. Dr. Gray, Dr. Nordtvedt, and Dr. Crawford, through their thoughts and ideas have enlightened me on several occasions. Thanks to Ashton for doing everything required to maintain our existence in this department, providing us with assistance and support. My special thanks to Daniel for making it very easy for us to get whatever we wanted – beginning from our daily cup of coffee, to all reference books for teaching, to registering us for courses, to providing support in the form of stationery and electronic equipment like desktops and photocopying. And for being a great friend, a person to talk to when we were exhausted.

My acknowledgements for the department are incomplete without mentioning the student assistants who have helped me with several requirements for my research. Notably, Yugesha Sapte whose role was very critical in creating a Python code that helped me extract all the relevant annual reports from annualreports.com. This drastically reduced my processing and data collection time. She was also instrumental in providing me support with some manual coding of my data. Also equally important was Sanjana Suresh who helped me with creating a Python script to extract keywords from all the annual reports and create an extensive dataset of all the key-terms related to digital transformation that occurred in those annual reports. In fact, this was the most important variable for me in my dissertation, my custom-created dependent variable. Acknowledgements are also due to many other student assistants (Laxmi, Rohit, and Poorvi) who helped me with some data related work from time to time.

Finally, the University's Office of Graduate Services, the College of Business, all the administrative officers and support staff, and The Graduate Studies Office for their Summer Dissertation Fellowship. Becky Nielson and Jairo for coordinating with us during our comps exams, without which I would not have been able to write my dissertation. William for providing all the administrative support. He has been the most proactive PhD coordinator I have experienced in the past 4 years. Ruthie Brock for helping me identify and procure the relevant datasets for me. All the other unnamed administrative assistants who must have worked behind the scenes to make my stay in the PhD program as comfortable as possible. Lastly, all the unnamed cleaning and janitorial staff who kept all our surroundings clean and hygienic enough for me to spend long hours in the office and work on finishing my dissertation.

DEDICATIONS

I offer this dissertation, this culmination of my four years of hard work to the One who has always been a source of great support – My Sadguru Shri Sai Baba. All through this difficult journey, I have sensed His Divine presence around me – which motivated me to take this work to its logical conclusion. He is like the Benevolent, Loving, Affectionate Mother who gently runs Her fingers through my hair and cheers me up in an instant. His spiritual presence compensates for the absence of physical proximity to my mother and father, who sit there, thousands of miles away from me, but always think of me, worry about me and my immediate family. I dedicate this work to my mother and father who shaped me to what I am today. During these 4 years, I was very grateful that my parents and especially my mother stayed with us here for several months to help our family through settling down and taking care of kids when my wife and I both had very grueling work schedules. I am also deeply grateful to my mother-in-law, who too made several long trips to the US to be with us, especially to take care of kids. Her presence too had provided us with great succor and relief in the midst of some very busy schedules.

Unquestionably, the biggest part of my dedication is dedicated to my wife Megha, because of whom I found love. The first love of my life, my life partner, my lifetime boss, my lifeline, has been an integral part all the trials and tribulations we faced during these past few years. We jointly faced some of our life's most challenging circumstances, few of which we are yet to overcome. But her enduring composure in the most difficult situations has kept me strong, kept me going. With her by my side for the rest of my life, I have the confidence that I will be able to achieve some great milestones. I pray that I am able to be with her for the rest of my life and beyond. As I write this dedication, I realize how much of a "taken-for-granted" attitude we display towards our loved ones. Our sweet kids, Arnav and Neerzaraa have enriched our lives

and provided us with the motivation to continue doing what I am doing. During these 5 years in this foreign country, both have had to endure cultural adjustments and frustrations. Both have had to be away from their close friends and other dear family members. Yet, they have soldiered on and are eventually getting adapted to the new way of life that beckons them. I dedicate the fruits of my hard work to both of them, who had to sacrifice so much to ensure that I pursued and completed my studies. With my upcoming life as an Assistant Professor, I hope I am able to compensate at least a part of all that they had to give up during last 4-5 years.

I also dedicate this work to some of my favorite past teachers, bosses, colleagues, and friends who encouraged me to embark on this difficult journey, which after 4 grueling years I have realized, has just begun!

TABLE OF CONTENTS

TABLE OF FIGURES

Digital transformation is not about technology….it is about strategy, leadership, and new ways of thinking.

- David Rogers (The digital transformation playbook)

CHAPTER 1 : INTRODUCTION

Businesses need to change with changing times. As tastes and technologies evolve, businesses need to remain relevant to their marketplace. At one point, Kodak dominated the market for photographic film such that an estimated 90% of all film loaded in cameras was Kodak. In 2012 this 120-year old company was forced into bankruptcy by its failure to commercialize the digital camera. In the near future the digital camera companies who put Kodak out of business may also be obsolete as people move to phones for photography. This is the age of rapid technological advances across a wide range of industries that have brought about large-scale disruptive changes .

The latest revolution to impact today's firms, popularly called "digital transformation", is broadly understood as "several digital innovations being brought together (Hinings, Gegenhuber, & Greenwood, 2018) to transform existing rules of engagement for and among organizations" or alternately the "use of technology to radically improve performance or reach of enterprise" (Westerman, Bonnet, & McAfee, 2014b). Digital transformation includes several related technologies including computing capabilities, networking, and data-storage and retrieval. Beginning with ERP (Enterprise-wide resource planning) and business intelligence, today's and tomorrow's technologies like BDA (Big Data Analytics), AI (Artificial Intelligence), ML (Machine Learning), IoT (Internet of Things), RPA (Robotic Process Automation), and Social

Media are revolutionizing many industries. The potential impact of this revolution is so pervasive, that it is being hailed as the fourth industrial revolution (Piccarozzi, Aquilani, & Gatti, 2018).

This fourth industrial revolution is based on automated and intelligent systems capable of communicating autonomously with an integrated command and control system across organizations (Piccarozzi et al., 2018). According to McKinsey Global Institute (Manyika et al., 2013)(Manyika, Chui, Buguin, Dobbs, Bisson & Marrs, 2013), the fourth revolution is the age of "cyber-physical systems" –that integrate computation, networking and physical processes, and include a myriad of technologies spanning mobile devices, IoT, AI, robotics, cyber-security and 3D-printing (Piccarozzi et al., 2018).

In general digital transformation as a technological revolution presents significant environmental challenges to contemporary businesses which will either disrupt current competitive advantage or provide a source of sustainable competitive advantage in the future (King, Grover, & Hufnagel, 1989; Mata, Fuerst, & Barney, 1995; Piccoli & Ives, 2005).

Today we see many large organizations mobilizing to adopt and take advantage of these technologies while others have struggled to keep pace. This raises important questions about how corporate leaders respond to rapid technological change. This dissertation proposes that the ability to adapt to digital transformation is primarily a result of managerial responses to these changes in the environment (Helfat & Martin, 2015; Rajagopalan & Spreitzer, 1997). This dissertation examines the micro-foundations of strategic change proxied by adoption of digital transformation among the S&P 500 from the period of 2008-2017. Specifically I test the role of CEO and top-management team characteristics in predicting digital transformation across different industries.

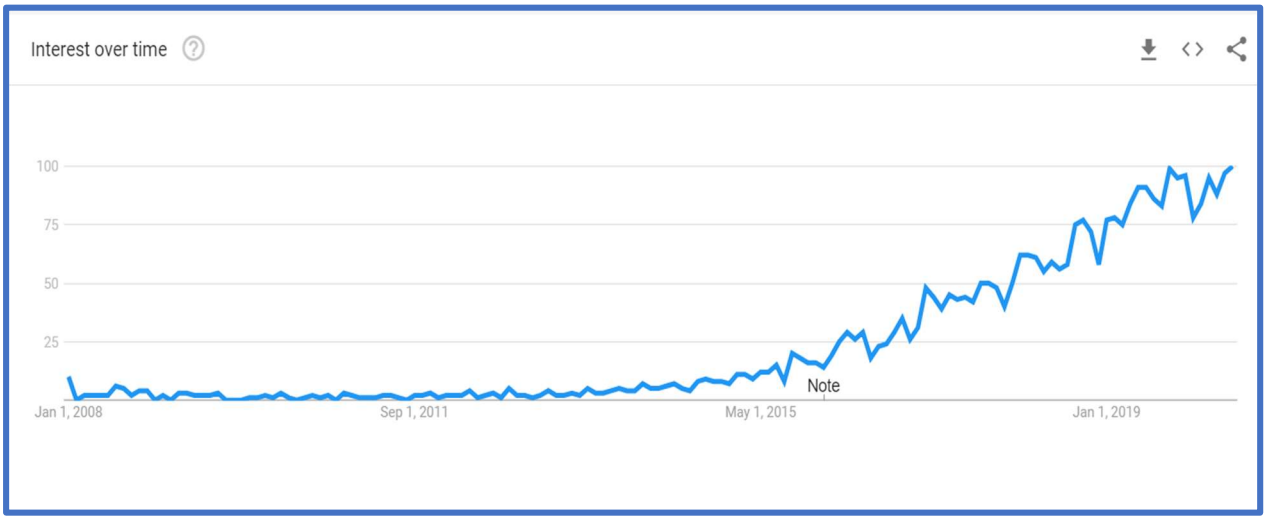
Due to the size and scope of digital transformation this phenomenon provides a unique opportunity to understand change at the strategic level. My dissertation examines the predictors of digital transformation as a unique window into strategic change itself. An organization undergoing digital transformation means that it is undergoing strategic change. Digital transformation fits the common definition of strategic change as both a phenomenon and as series of acts of innovation (Crossan & Apaydin, 2010). The choice by organizational decision-makers to incorporate any change of strategic nature might lie anywhere between the continuum of deliberate to emergent strategies (Mintzberg & Waters, 1985). In other words, whether a strategy or change is intended (deliberate) or not (emergent), this choice is made by the organizational decision-makers.

This dissertation also focuses on the characteristic of CEOs and top managers as predictors of strategic change because digital transformation requires significant investment, process change, and risk. I contribute to the large volume of research which examines upper echelon characteristics and different aspects of strategic change by examining the role that CEO and TMT characteristics play in driving strategic change in stable environments versus highly dynamic environments. CEOs and top managers have better visibility, clout and wherewithal to make choices, take decisions and execute those decisions in all organizations. This dissertation contributes to this literature by developing theory that the magnitude of the effect of upper echelon characteristics on strategic change is contingent upon the extent of dynamic context. Under certain conditions of environmental dynamism, upper echelon characteristics strongly influence strategic change decisions whereas under other conditions, the effect is weak or non-existent.

CHAPTER 2: LITERATURE REVIEW

Corporations today are increasingly discussing and implementing digital transformation. For instance, a worldwide Google Trends search on the term "digital transformation" appearing in "business and industrial" category of google search showed a value of 2 in July 2012 on the trends index, and a value of 100 in February 2020, i.e. a rise of 5000% in 7 years[1]. Please refer to figure 1 for this graph below.

Figure 1: Google Trends as of May 2020 (Source – Google Trends)



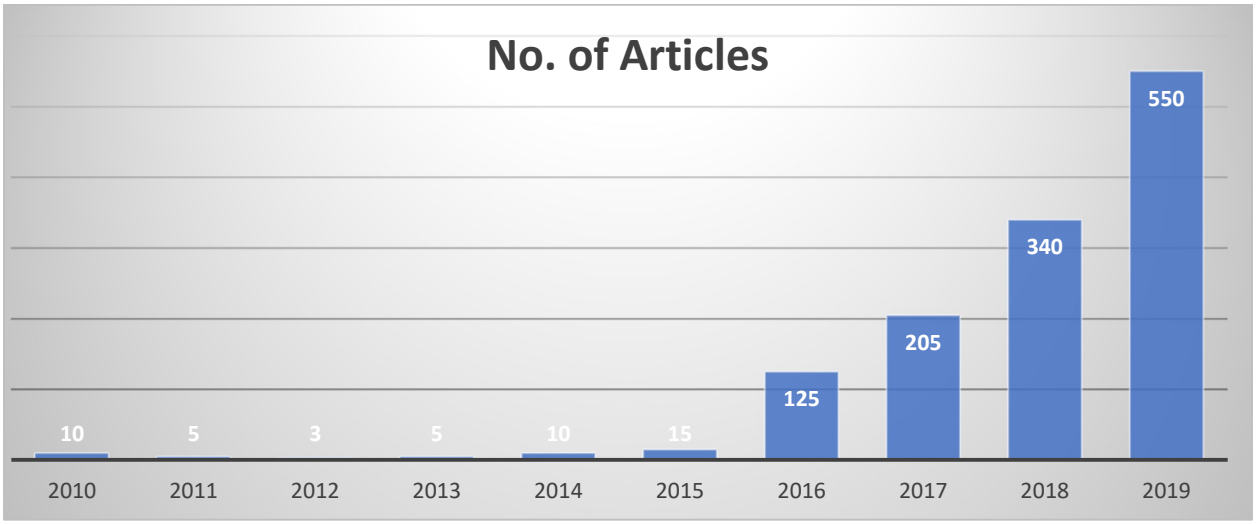
In academic literature research related to "digital transformation" has grown from just 3 to 5 articles in the years 2006-2008 to 680 articles in 2019.[2] This indicates the growing relevance

[1] https://trends.google.com/trends/explore?cat=12&date=all&q=%2Fm%2F0g5r88p

[2] http://wcs.webofknowledge.com.ezproxy.uta.edu/RA/analyze.do?product=WOS&SID=5FqoNInR6whrbypK132&field=PY_PublicationYear_PublicationYear_en&yearSort=true.

of digital transformation as a phenomenon and an example of strategic change within organizations. Please refer to figure 2 for this graph.

Figure 2: Increasing trend of academic articles on digital transformation (Source – Web of Science)



NATURE OF STRATEGIC CHANGE

Digital transformation provides a unique opportunity to study a specific strategic change initiative embarked upon by organizations. Strategic change is defined by Rajagopalan & Spreitzer (1997, p. 49) as, "a difference in form, quality or state over time in an organization's alignment with its external environment" which is based on work by Van de Ven and Poole (van de Ven & Poole, 1995). (Chaffee, 1985) suggests that organizations make strategic choices to undergo change in response to their changing environment. These choices are typically exercised by the executives or managers of the organization (Andrews, 1971; Boeker, 1997; Child, 1972; Schendel & Hofer, 1979). Finally, Carpenter et al. (Carpenter, Geletkanycz, & Sanders, 2004) define these outcomes as "strategic outcomes" that include business, corporate, international,

change, strategic interactions, and policies, which in turn lead to “performance outcomes” such as financial returns, market performance, social outcomes and innovation.

Strategic Choices and Outcomes, Strategic Change and Organizational Change

Within management literature, macro scholars mostly address change as “strategic change” whereas micro-related scholars mostly refer to change as “organizational change”. Even as phenomena and actions related to change may be same, macro scholars study firm or organization-level outcomes as units of analysis whereas micro scholars study individual or team level outcomes within organizations. This dissertation addresses the micro-foundations of strategic change and therefore includes both the micro and macro perspectives.

Strategy is as multi-faceted and as complex as organizations (Chaffee, 1985). To simplify and provide a structure to such complexity, she suggests that literature related to strategy can be classified into three different models, linear, adaptive and interpretive. Chandler’s classical interpretation of strategy represents the linear model, defined as “determination of the basic long-term goals of an enterprise, and the adoption of courses of action and the allocation of resources necessary for carrying out these goals” (Chandler, 1962), p.13). Hofer (Hofer, 1973) is an example of the adaptive model definition, “strategy is concerned with the development of a viable match between the opportunities and risks present in the external environment and the organization's capabilities and resources for exploiting these opportunities” (p 3). For the interpretive model, Chaffee defines strategy as “orienting metaphors or frames of reference that allow the organization and its environment to be understood by organizational stake-holders, motivating them to believe and to act in ways that are expected to produce favorable results for

the organization" (Chaffee, 1985, p 93). Taken together these views suggest that strategy includes three main elements: (a) existence of long-term goals for the firm or organization, (b) availability and allocation of resources by organizations to address the opportunities and risks due to environmental changes as understood by stake-holders, (c) belief by organizational stake-holders that their decisions and actions will lead to favorable results.

Strategic change is defined as "an empirical observation of difference in form, quality, or state over time in an organizational entity" (van de Ven & Poole, 1995). These differences over time have been examined in the literature from three primary perspectives: context, content, and process (Armenakis & Bedeian, 1999; Rajagopalan & Spreitzer, 1997). Context focuses on the conditions governing the organization's internal and external environments, a combination of answers to three questions – "why", "where", and "when" of strategic change. Content focuses on the substance, or the "what" part of strategic change. Finally, process explains the actions taken during change efforts, or the "how" part. In other words, the answer to the question "why, where and when does strategic change occur?" is the context, "what comprises of strategic change?" is the content and "how does this strategic change occur?" is the process.

Context. The context for strategic change is defined by the organization's environment, internal and external. Internal factors like organizational size, age and inertia determine the organization's effectiveness in responding changes in the external environment (e.g. regulation changes, competition) (Armenakis & Bedeian, 1999). Context includes industry forces like current or prospective competitors, stake-holder behavior and product/ service alternatives (Porter, 1980). The change context (comprising of external environmental conditions and internal organizational conditions) addressed by managerial cognitions and actions (process) changes to strategy and thus organizational outcomes (Rajagopalan & Spreitzer, 1997).

Content. Content of strategic change is characterized by change in its scope, resource deployments, competitive advantages, and synergy (Hofer & Schendel, 1978; Rajagopalan & Spreitzer, 1997). A well thought out strategy empowers a firm to face competition from its rivals in the contextual space (Hofer & Schendel, 1978; Thorelli, 1977), This means that better the conceptualized strategy, better are the chances of the firm being able to create its competitive advantage. The content of strategic change manifests itself at two different levels, business and corporate. Business level strategy is when an organization makes changes in its strategy content within a specific business; for example changes related to strategies in market entry, new product development, product range diversification, competitive actions, and technological upgradations. When the organization changes its mix of different businesses through mergers, acquisitions, alliances, spin-offs or divestitures, it may be called strategic change at corporate level.

Process. Change in strategy process includes the changes in activities leading upto and supporting the choice of strategy (Huff & Reger, 1987). The process view of strategy looks at stakeholder actions and responses in executing strategic change initiatives. These actions can occur at the level of the external environment or the firm or the individual (Armenakis & Bedeian, 1999). Several models of change process have been proposed over the years beginning with Lewin (Judson, 1991; Lewin, 1947). Most commonly used today is the 8-step model detailed by Kotter (Kotter, 1995). Armenakis & Bedeian (1999, p. 301) describe this model in detail. :

(a) establishing a sense of urgency by relating external environmental realities to real and potential crises and opportunities facing an organization, (b) forming a powerful coalition of individuals who embrace the need for change and who can rally others to support the effort; (c) creating a vision to accomplish the desired end-result; (d) communicating the vision through

numerous communication channels; (e) empowering others to act on the vision by changing structures, systems, policies, and procedures in ways that will facilitate implementation; (f) planning for and creating short-term wins by publicizing success, thereby building momentum for continued change; (g) consolidating improvements and changing other structures, systems, procedures, and policies that aren't consistent with the vision; and (h) institutionalizing the new approaches by publicizing the connection between the change effort and organizational success.

Multi-lens framework theory of strategic change

This content and process reflects an upper echelons perspective of organizations with the roles and responsibilities of senior leaders in driving strategic change. This perspective is best described by Rajagopalan & Spreitzer's (1997) multi-lens framework which details how managerial cognitions get converted to managerial actions in pursuit of strategic change. To create a comprehensive model of strategic change, Rajagopalan and Spreitzer (1997) reviewed and integrated several streams of literature. They classify various theory and empirical studies related to strategic change can be classified into three categories.

Rational-lens perspective. The rational-lens perspective is : Past literature about strategic change in viewed as a discrete phenomenon in which change is considered or studied as a unitary concept, can be classified in this category. Here, unitary concept means that only the content of strategy changes, not the organization or environment. Such change is measured either as likelihood, magnitude, direction, or timing of change. For example, Goodstein and colleagues (Goodstein, Gautam, & Boeker, 1994) studied effects of board size and diversity on strategic change in the form of change in scope and organization of services provided. In general But this

approach of a rational lens disregards managerial actions and cognitions, treating them as a black box in predicting outcomes., resulting in conflicting findings in several studies. It is well-established that performance is impacted not only by strategy content, but also the change in organizational or environmental context. This does not get captured in a rational-lens model because there is an inherent assumption that such contextual changes are deterministic (happen automatically rather than being driven). This is a shortcoming as performance is impacted not only by strategy content, but also the change in organizational or environmental context. To some extent, this drawback of rational lens model is addressed in the learning lens model.

Learning-lens perspective. The learning-lens model of change assumes the interaction of content of strategy and organizational or environmental conditions. Through this lens, strategic change is characterized in two different ways. First is that change that is an ongoing gradual or continuous evolutionary process. Second is that change is more often sudden, discontinuous and revolutionary or transformational. Learning-lens perspective views strategic change as a combination of change in content of strategy and change in organizational or environmental context. It is often described as an iterative process which managers take to probe the environmental (external) and organizational (internal) context and, learning through each step. Managers then make decisions about change (whether to resist or initiate) by iteratively interacting with and influencing the context and being influenced by it. Managers may study impending or occurred change in external environmental context, like advent of new technologies or environmental volatility, to determine how to proceed in their own organizations. Similarly, managers may also drive changes due to changes in internal context of organizations, like changes in leadership or drop in performance. In either of these cases, managers may take either proactive or reactive actions to reduce the uncertainty. Even as external and internal

contexts and managerial actions are clearly identified in their learning-lens perspective, Rajagopalan and Spreitzer have observed that some past studies were empirically not able to establish a causal relationship. For this reason, they suggest an additional lens which they call as a "cognitive" lens perspective.

Cognitive-lens perspective. A cognitive perspective to strategic change recognizes the role of managerial cognitions in the change process in addition to the external (environmental) and internal (organizational) contexts. Managerial cognitions means knowledge structures, core beliefs, cause maps and schemas (Walsh, 1995). The cognitive lens perspective creates a framework to explain the origin of the decisions and subsequent actions. This lens is able to causally link the managerial characteristics to the strategic actions that they take. Thus, the actions of managers act as operationalized proxies for measuring the changes in content of strategy. A key aspect of cognitive lens is that environment cannot be objectively determined; it is a perceptual construct in the mental maps of the managers. Despite the refinement over rational and learning lens models, cognitive lens too suffers from some drawbacks. Rajagopalan and Spreitzer say that it may be unable to clearly distinguish managerial cognitions and actions from actual changes in the content of strategy, hence propose a multi-lens model.

Multi-lens framework: The multi-lens framework integrates all the three perspectives – rational, learning, and cognitive. This incorporates several theoretical aspects of strategic organizational change. The rational perspective explains the deterministic approach towards change as proposed by population ecology theorists. The learning perspective explains the iterative processes that go into evolutionary changes; similar to contingency theorists' approach to strategic actions. But without the cognitive lens approach, antecedents of such strategic actions cannot be explained. Also, these cognitive abilities will determine the extent of

willingness or readiness to drastically change underlying knowledge structures as in revolutionary changes. Managers use their cognitive abilities to make decisions about environmental and organizational factors to make appropriate strategy changes.

The multi-lens framework also integrates the levels of analysis typical of studies on change. The rational lens perspective by itself can only study impact of environmental phenomena at the firm level. But with learning and cognitive lens perspectives, individual managers can be studied. According to the Rajagopalan and Spreitzer (1997) model, environmental conditions and changes, and organizational conditions and changes affect managerial cognitions, which in turn will drive strategic decisions. A similar conceptualization is also presented by Zajac et al (Zajac, Kraatz, & Bresser, 2000) while explaining strategic change and dynamic strategic fit. They use the terms environmental contingencies (similar to external environment) and organizational contingencies (internal environment) that influence the desirability of strategic change. In other words, change is driven by managers in response to contextual conditions. Theory for this dissertation is developed using the assumptions of the multi-lens framework (Rajagopalan & Spreitzer, 1997) in and the interaction between managerial cognition based on upper echelon theory and the external competitive environment.

UPPER ECHELONS PERPECTIVE

To examine the role of managerial cognition this dissertation uses upper echelons theory to predict that managerial characteristics in part drive organizations' strategic choices towards digital transformation. The role of senior organizational leaders or upper echelons in driving organizational outcomes is a basic tenet of strategic management research (Hambrick & Mason,

1984). Vast numbers of empirical studies have used executive characteristics, either individually, or collectively, as proxies for strategic choice or decision-making in organizations (Bromiley & Rau, 2016). The term "strategic choice" is borrowed by Hambrick and Mason from Child (1972). It is defined as, "a fairly comprehensive term to include choices made formally and informally, indecision as well as decision, major administrative choices (e.g., reward systems and structure) as well as the domain and competitive choices more generally associated with the term 'strategy" (Hambrick & Mason, 1984), p. 195). This describes the notion that organizational outcomes are partly dependent on the managerial characteristics of the top management (Hambrick & Mason, 1984) and strategic choices by these individuals and groups are in turn driven by their individual characteristics. This link between observable characteristics of top managers, strategic choices and firm performance is called Upper Echelons theory (Carpenter, Geletkanyz, & Sanders, 2004; Hambrick, 2007; Hambrick & Mason, 1984).

Carpenter and colleagues (Carpenter et al., 2004) define three interconnected tenets on which this mechanism rests. One, strategic choices of organizations are a reflection of values and cognitive bases of the powerful coalition of actors that drive the organization. Second, inherent values and cognitive bases are not directly observable but observable characteristics like education and work experience may represent such values and bases. Third, observable characteristics may influence significant organizational outcomes.

Based on these tenets, Hambrick further explains this mechanism in three steps (Hambrick, 2007). Firstly, executives' perspectives based on their values, personalities, and experiences affect their field of vision; in other words – the directions they look and listen. Next, this leads them towards selective perception – what they actually see and hear. In the third step,

these executives attach meaning to what they see and hear based on their attributes and interpret accordingly.

Upper echelons theory presumes that strategic choices are influenced by perceptions and information processing abilities of decision-makers. Research based on upper echelons relates to top executives of firms, mostly boards, CEOs, or TMTs (top management teams) or their combination. Bromiley & Rau (Bromiley & Rau, 2016) have extended this understanding further in the context of strategy process, where strategy process is defined as "the mechanisms by which organizations formulate and implement strategy" (p 174). They have classified all the literature related to upper echelons into two broad categories, studies examining cognitive influences on executives and studies examining social/ behavioral influences on executives. Research on cognitive influences considers the "cognitive base" including the attention, perception, cognition and information-processing abilities of such CEOs and top managers (Bromiley & Rau, 2016). Research on social/ behavioral influences include demographic and human capital characteristics of top managerial personnel. Common executive characteristics studied using upper echelons theory include tenure, experience, and gender. Another set include personality, values and affect. The third set include group characteristics such as experience, roles and social ties. In my study I examine some of these characteristics like mean tenure and role heterogeneity, hence I elaborate on the mechanisms here.

Every individual, depending on his/ her inherent values and "cognitive base" has a perspective different from others. This unique perspective is a natural outcome of the quantity and quality of information and experience that an individual processes. The manner in which this happens is known as "cognitive style" (Messick, 1976) that drives perceiving and judging information (Hough & Ogilvie, 2005). Due to their different cognitive styles, executives frame

their strategic options, decisions and actions based on their particular perspective. The information processing from the time they are faced with a situation to their actual decision is governed by "bounded rationality" (Cyert & March, 1963; March & Simon, 1958). Bounded rationality is defined by Simon as "rational choice that takes into account the cognitive limitations of the decision-maker - limitations of both knowledge and computational capacity" (Lipman, 1995; Palgrave, 1987). Similarly, strategic choice can be explained as a rational choice under cognitive limitations.

In general, top managers are overwhelmed by stimuli much larger than they can cognitively handle (Cohen & Levinthal, 1990). This limits their ability to process the available information resulting in paying focused but limited attention (Hambrick & Finkelstein, 1987). To cope up with such information- overload especially in rapidly changing environments, managers use mental models or interpretive schemas of "noticing and constructing meaning" (Barr, Stimpert, & Huff, 1992). These schema and heuristics in turn are partly driven by individual managerial characteristics. There could be several stimuli to trigger information overload. One such stimulus, "job demand" has been proposed by Hambrick and others (Hambrick, Finkelstein, & Mooney, 2005) as possible explanation for executives to increasingly rely on their individual characteristics. They define executive job demand as "the degree to which a given executive experiences his or her job as difficult or challenging" (p 473). Research supports job demand as a critical moderator for the relationship between top management characteristics and strategy. Goll, Brown-Johnson & Rasheed (Goll, Brown Johnson, & Rasheed, 2007) studied the US airline industry and found that managerial characteristics predicted strategic choices more strongly in environments that created greater job demands. Another study of airline industry by Cho and Hambrick (Cho & Hambrick, 2006) mirror these findings that deregulation of the

industry partly caused managerial attention to be focused toward the change in environment which in turn influenced strategic change.

While it is clear that values, cognitive bases and perceptions of top executives have a direct relationship with executive decision-making, these characteristics are not always directly measurable due to unavailability or reluctance of such senior managerial personnel to subject themselves to psychological evaluation (Hambrick & Mason, 1984). To overcome this inadequacy, upper echelons theory suggests that observable characteristics can be used as reasonable proxies for such underlying values, cognitions, and perceptions (Carpenter et al., 2004). Bromiley & Rau (2015) have classified upper echelons characteristics as observable, underlying, and interaction with others. Observable characteristics include age, functional background, career experiences, education, socio-economic roots, and financial position. Underlying psychological characteristics include personality, core self-evaluation, charisma, humility, narcissism, hubris, overconfidence, values, affect, and intelligence.

Past literature has studied other CEO characteristics like age, gender, education, firm tenure or firm experience, career experience, CEO attitude towards change, CEO innovativeness, among several others. These have been used over time to study a number of different strategic choices such as foreign market entry mode (Herrmann & Datta, 2006), likelihood of environmental disclosure (Lewis, Walls, & Dowell, 2014), and firm strategic persistence (Datta, Rajagopalan, & Zhang, 2003). CEO decisions and actions examined include firm R & D spending (Barker & Mueller, 2002), information technology adoption (Abdul Hameed & Counsell, 2012; Thong & Yap, 1995), and innovativeness (Kitchell, 1997; Lefebvre & Lefebvre, 1992; Lin, Lin, Song, & Li, 2011). Impact of CEO characteristics on organizational outcomes have included corporate social performance, corporate sustainable development, quality of

internal control, and organizational culture; apart from the usual financial parameters like RoA, RoS and MTB (e.g. Quigley & Hambrick, 2015). This dissertation examines individual CEO characteristics like total firm tenure and CEO power as predictors of digital transformation.

Upper echelons theory can also be used to examine the characteristics of top-management means using the same assumptions. Collectively individual characteristics combined together (e.g. taking average of age, or tenure, or years of experience, or human capital, etc.). Bromiley & Rau's interaction or combination characteristics may all be classified under the category "composite" characteristics. This fits with human capital theory that conceptualizes aggregated individual characteristics like experience, education and skills as resources for teams and top management teams in particular (Penrose, 1959; Rumelt, 1984; Schultz, 1971; Teece, 1982; Wernerfelt, 1984) that can be leveraged to provide sustainable competitive advantage to organizations (Barney, 1991; Peteraf, 1993; Ployhart & Moliterno, 2011).

Becker (1962) refers to human capital as the learned skills and knowledge that individuals develop through their prior experience, training, and education (Helfat & Martin, 2015). Certain other conceptualizations also incorporate psychological attributes like cognitive ability, personality, values, and interests while discussing aggregated human capital at team or organizational levels (e.g. Ployhart & Moliterno, 2011). In this dissertation, I focus on human capital resources of the top management teams. As these are a part of upper echelons, the highest decision-making body of the organization, the human capital resources of these teams may be considered as a proxy for the strategic human capital resource for the organization. Top executives or upper echelons, through their knowledge and skills thus acquired, can use sense-making process to detect opportunities and threats, and then use sense-giving to drive the strategic change initiatives (Gioia & Chittipeddi, 1991).

Upper-echelons theory has been used to examine influence of TMT characteristics on strategic decisions and strategic change. This is an extension of the notion of dominant coalition as proposed by Cyert & March (1963). The idea of dominant coalition represented by the senior executives of the organization influences strategic direction and organizational performance. At a macro level, this group of senior members is considered the interface between the organization's external and internal environment. Being at the top hierarchical level their choices, decisions, and actions are considered powerful enough to have an impact on the whole organization. Many previous studies have addressed the collective and composite influence of such top management teams. Some of the collective characteristics studied in the past are team size (Certo, Lester, Dalton, & Dalton, 2006; West & Anderson, 1996) and team openness (Amason & Sapienza, 1997).

Some of the composite characteristics studied in the past are team heterogeneity (Hambrick, Cho, & Chen, 1996; West Jr & Schwenk, 1996), team group processes (West and Anderson, 1996) and team diversity (Knight, Pearce, Smith, Olian, Sims, Smith, & Flood, 1999). For example, Hambrick, Cho and Chen (1996) found that team heterogeneity was positively associated with the propensity to take action, boldness of the action as well as magnitude of the competitive actions and negatively associated with the speed of the actions executed. Some outcomes studied with respect to TMT characteristics are organizational performance in market share and profits (West Jr. and Schwenk, 1996), innovation (West and Anderson, 1996), cognitive and affective conflicts (Amason & Sapienza, 1997), strategic consensus (Knight et al., 1999)

ROLE OF THE ENVIRONMENT

Change in organizational strategy is explained in literature using two divergent perspectives mapped on a continuum (Boeker, 1997; Gersick, 1994) i.e. from inertial to adaptive. The inertial perspective emphasizes the view that organizations prefer status-quo. Their ability to adapt to changing environment is constrained by several factors like structural rigidity, political resistance and vested interests (Hannan & Freeman, 1984; Tushman & Romanelli, 1985). Such organizations may find it difficult to change. On the other end of the continuum is the adaptive perspective. This perspective emphasizes that managers or executives use mechanisms like sense-making and sense-giving (Gioia & Chittipedi, 1991) to track environmental changes and thus initiate changes to their strategies and actions within their organizations (Boeker, 1997; Child & Mansfield, 1972)(Boeker, 1997; Child & Mansfield, 1972)(Boeker, 1997; Child & Mansfield, 1972).

Both perspectives, inertial and adaptive, have merit and different scholars have found that either of the perspectives is predominant under certain conditions. For example, Kelly and Amburgey(Kelly & Amburgey, 1991) found perceived favorability of environmental change, organizational age, and prior change experience moderated the relationship between environmental change, organizational size and strategic corporate change. In other words, the inertial perspective was dominant only in conditions where perceived environmental change is unfavorable or when organizational age was higher or when organization had not experienced similar environmental turbulence earlier.

In support of the adaptive perspective, Boeker's (1991) study found that top managerial characteristics like average team tenure and diversity influenced strategic change independent of

the environmental conditions. Similarly, Gioia and Chittipedi (1991) observed that CEO and top management team drive strategic change through a 4-stage process called envisioning, signaling, revisioning and energizing as moderated by environmental conditions. Some of the moderators in the relationship between upper echelons characteristics and strategic change were environment (Goll, Brown Johnson, & Rasheed, 2007) or environmental dynamism (Jiao et al., 2013), CEO power (Haynes & Hillman, 2010), and industry characteristics (Datta et al., 2003).

The decision to make changes in strategy is a managerial choice based on environmental conditions, organizational factors, and managerial cognition. A fundamental notion in strategy and organizational literatures is the influence of environmental characteristics on the strategies and performance of firms (Hannan & Freeman, 1984). One form of "environmental conditions and changes" as explained in the Rajagopalan & Spreitzer model of strategic change, is "environmental volatility" or "environmental dynamism", defined as the level/ rate of instability or unpredictability of change faced by an organizational unit (Barry, Kemerer, & Slaughter, 2006; Dess & Beard, 1984; Dugal & Gopalakrishnan, 2000). Empirical evidence suggests that in volatile environments, dynamic capabilities of managers are more effective in implementing organizational change (Jiao, Alon, Koo, & Cui, 2013).

The relationship between environmental dynamism and managers' strategic choices including organizational change is further elaborated in the dynamic capabilities literature (Helfat & Martin, 2015; Teece, 2007; Teece, Pisano & Schuen, 1997). "Dynamic" here means renewability of competences to match with demands of the changing environment. The term "Capabilities" "emphasizes the key role of strategic management in appropriately adapting, integrating, and reconfiguring internal and external organizational skills, resources, and functional competences to match the requirements of a changing environment" (Teece et al,

1997, p 515). The dynamic capabilities of managers are called “dynamic managerial capabilities” (Adner & Helfat, 2003). Helfat and Martin define these as “the capabilities with which managers create, extend, and modify the ways in which firms make a living—to help explain the relationship between managerial decisions and actions, strategic change, and corporate performance under conditions of change” (p: 1282). The top executives or top managers of firms thus try to respond to their environment to the best of their capabilities.

Managers of organizations effectively transform information and knowledge into innovative products, services, and processes with the help of dynamic capabilities like opportunity-sensing, reconfiguring, organizational flexibility and technological flexibility. Opportunity-sensing refers to deep understanding of market development opportunities. Reconfiguration capability means developing, configuring, integrating, innovating, and updating resources and operational processes. Technical flexibility capability alludes to capability of rapid improvement in existing technologies to meet customer needs. Organizational flexibility capability refers to organizational structure’s attributes concerning decision-making processes, task configuration and information flow. An example could be allowing decision-makers to break through formal procedures in order to maintain working flexibility and dynamism (Jiao et al., 2013).

Adner and Helfat identify three core “underpinnings” of dynamic capabilities that influence strategic change – managerial cognition, managerial social capital and managerial human capital. Helfat & Martin (2015) have classified managerial cognition as comprising of these elements – (a) mental models and beliefs or knowledge structures (Barr et al., 1992; Eggers & Kaplan, 2013; Walsh, 1995); (b) mental processes and managerial cognitive capabilities (Helfat & Peteraf, 2015); (c) emotions (Hodgkinson & Healey, 2011). These components of

managerial cognition influence the way biases and heuristics are formed while sensing market changes, understanding the implications of different choices, and ultimately taking action (Garbuio, King, & Lovallo, 2011). This aligns well with the philosophy of upper echelons, i.e. the manner in which cognitive base and values drive the limited vision, perception and interpretation of managers in determining strategic choices.

Each manager or executive or director may have a unique mix of knowledge and skills background in terms of functional and technological domain, and industry-specific and firm-specific expertise that governs these choices (Helfat & Martin, 2015). This in turn will drive differences in their respective absorptive capacity (Cohen & Levinthal, 1990) for sensing different types of stimuli. These individual differences are amplified and combined to organizational level through various "emergence-enabling processes" (Ployhart & Moliterno, 2011) that combine two inter-related components. One component is the unit's or organization's task environment ranging from simple to complex depending on level of inter-dependence required among the organizational or team members. This range of complexity of the task environment manifests the dynamism of the external and internal contexts (Eisenhardt & Martin, 2000). Another component according to Ployhart and Moliterno, is the organization's or team's "emergence-enabling states" that comprise of cognitive (think), affective (feel), and behavioral (act), and mechanisms (Kozlowski & Ilgen, 2006). Thus, the collective perception of organizational task environment complexity is contingent upon the way organizational members think, feel and act at individual levels. This process of aggregation and its consequences are explained through the dynamic capabilities perspective as driving complementarities among the group members, eventually leading to positive organizational outcomes (Wright, Coff, & Moliterno, 2014). This aligns with my earlier described classification of units of analysis for

upper echelons characteristics (individual, collective, combination and composite) as probable explanation for how and why certain upper echelons characteristics influence various organizational outcomes, especially change.

Antecedents and Moderators Of Strategic Change – Internally driven versus Externally driven?

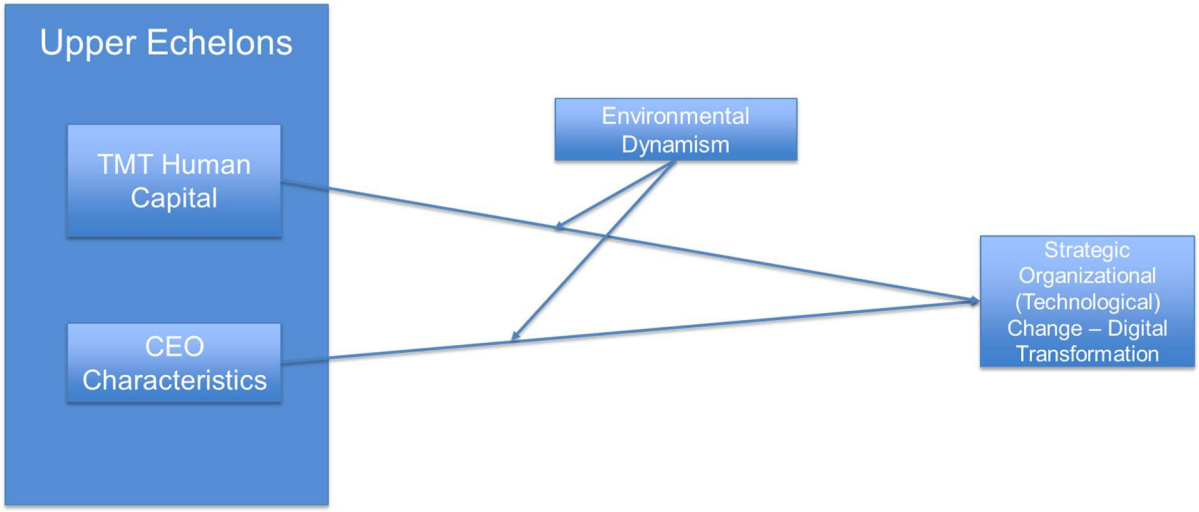
According to the Rajagopalan and Spreitzer (1997) model, environmental conditions and changes (i.e. external environment and dynamism), and organizational conditions and changes (i.e. internal environment and dynamism) affect managerial cognitions, which in turn will drive the strategic decisions taken by them. In other words, any change may either be driven by external environmental conditions or by internal organizational conditions or by a combination of both. But in any of these contingencies, managerial cognition will determine the actions that are taken. I have explained earlier the process and constraints under which managerial cognition drives strategic choices. It would mean that given the constraints under which managerial cognition operates, the external environmental conditions or internal organizational conditions will moderate the decisions that managers take. It also further means that the size of the impact of managerial cognition on the strategic outcomes will depend on how intense or how strong these moderating conditions are. This is the basic premise on which my study is constructed. I elaborate this using the context that digital transformation is a strategic change for the firm. I also explain environmental dynamism as the external environmental context, all of which I hypothesize and test in the following chapter.

CHAPTER 3: HYPOTHESES DEVELOPMENT

As described in the theory development section, past research has already demonstrated that strategic decision-making can be predicted by a number of executive characteristics, either individually, or collectively, or in combination by multi-dimension scales, or as composites. The decision-making processes may comprise of elements like intent and propensity which will lead to strategic choices, strategic actions, and organizational outcomes. These in turn, eventually influence organizational performance. Also, as explained earlier, digital transformation - the fourth industrial revolution, is an appropriate representation of strategic change due to changes in the external environment. The intent, propensity, and magnitude of making a choice about initiating and driving digital transformation can thus be driven by a combination of these executive characteristics with industry pressures for such a transformation. In other words, executive characteristics in combination with the extent of environmental influence of digital transformation on an industry will determine the extent of adoption of similar measures within a firm. Also as described earlier, past literature has studied TMT characteristics in some studies and CEO characteristics in some studies. But an attempt to segregate CEO characteristics from the other TMT members has been very rare. I thus create my hypotheses to incorporate all these intricacies.

Therefore, I examine the predictors and outcomes of strategic change using CEO and TMT characteristics to predict adoption of digital transformation. For this, I use a sample of publicly traded firms that constitute the S & P 500 Index. My full model is presented in Figure 3 below.

Figure 3: Full Study Model



I develop a series of hypotheses that predict the ways in which these relationships can be tested.

TMT Human Capital and Digital Transformation

TMT characteristics in general including tenure and experience have been found to be associated with strategic change (Goll et al, 2007; Wiersema & Bantel, 1992). For example Wiersema and Bantel (Wiersema & Bantel, 1992) found shorter organizational tenure, higher team tenure, higher mean team education level, and higher education specialization heterogeneity were all positively related to strategic change measured as diversification. In another study, Cho and Hambrick (2006) found TMT demographics to be related to strategic change as measured by shift of orientation (Cho & Hambrick, 2006). They found that managerial attention moderated this relationship.

TMT experience should also predict strategic change as measured by digital transformation. Each member of a TMT has different roles and in many cases, the TMT member

may have performed multiple roles in the firm. This could be conceived as role homogeneity or heterogeneity, depending on the number of roles held in the firm. As each TMT member represents a different functional role, chances of homogeneity are very low, but still existent. This homogeneity may be limited to industry experience or organizational tenure. To that extent, I predict that experience heterogeneity will be positively related to digital transformation.

The reason is that role heterogeneity provides broader experience for TMT decision-making and information processing based on the "cognitive bases" and values of the individual (Carpenter et al., 2004; Hambrick, 2007; Mischel, 1977). Collective experience affect the cognitive maps or heuristics Barr (Barr et al., 1992) used by managers. For example, the team level is the study by Wei and Wu (Wei & Wu, 2013) found that cognitive diversity among TMT members was positively associated with elaboration of task-information, which in turn was associated with better financial and innovation performance. (Ferrier, 2001) found that top management team heterogeneity was related to competitive aggressiveness. Heterogeneous teams were positively associated with higher complexity of attack and negatively associated with duration of a competitive attack. This work illustrates how different cognitive styles (Hough & Ogilvie, 2005) and cognitive stimuli (Goll et al., 2007) influence decision-making process of executives. Therefore my first hypothesis:

Hypothesis 1(a): TMT Role heterogeneity has a positive relationship with digital transformation

However, this relationship also depends on environmental conditions. In the inertial perspective of environmental influence (Hannan & Freeman, 1984), the environment directly influences strategy or strategic change (rational lens model of Rajagopalan – Spreitzer framework). But the adaptive perspective (Boeker, 1997) shows that top managers of the organization use their capabilities to choose whether to change or not and to what extent. So, when these top managers sense environmental dynamism, they will take conscious decisions on what to do next. However, as these managers use their cognitive capabilities to sense, their choices will be a reflection of their past experiences. If in the past they have experienced high dynamism, they will be familiar with or comfortable with such dynamism. Therefore, their actions will reflect strategic actions leading to change. However, if the top managers have had a reasonably stable past experience, then they will be unsure of what to do next because of the uncertainty, leading to an inertial situation where the top executives may not initiate changes to status quo. Therefore, environmental dynamism moderates the relationship between TMT role heterogeneity and digital transformation.

Hypothesis 1(b): The relationship between TMT role heterogeneity and digital transformation is moderated by environmental dynamism such that in less dynamic context the relationship will be negative whereas in highly dynamic context the relationship will be more positive

TMT experience can also be measured as team tenure. Top management teams have to work in unison for making appropriate choices, decisions, and actions. So, the more time they

spend together, chances are that their team-coordination and synchronization will be higher. This is evident in some past studies which found that higher average or aggregate level of certain depth of characteristics (e.g. team tenure) were positively associated with strategic change (Wiersema & Bantel, 1992). As time spent with each other goes on increasing, initial issues related to trust and understanding will reduce, thus increasing the cooperation and coordination among the TMT members. But as time together goes on increasing, it might generate several issues like interpersonal conflicts, ego clashes and inter-departmental/ inter-functional politics (Tushman & Romanelli, 1985). In a paper which studied the relationship between tenure and environmental context (Keck, 1997), the author found that shorter tenured, heterogeneous senior executive teams addressed environmental turbulence more effectively. Thus, depending on the time spent together, different outcome scenarios may arise. So, the next hypothesis:

Hypothesis 2(a): There is an inverse u-shaped curvilinear relationship between TMT mean tenure and digital transformation, such that digital transformation will be lowest at extremely low and high mean tenures

Average organizational tenure is negatively related (Finkelstein & Hambrick, 1990) and average TMT tenure is positively related to (Wiersema & Bantel, 1992) propensity to change strategy. Due to these different findings, it is clear that there are some moderating conditions to the relationship between TMT tenure and strategic change. Higher mean tenures may lead to inertial tendencies owing to cognitive limitations of not being able to perceive the environment objectively. This may lead to a negative relationship between TMT mean organizational tenure

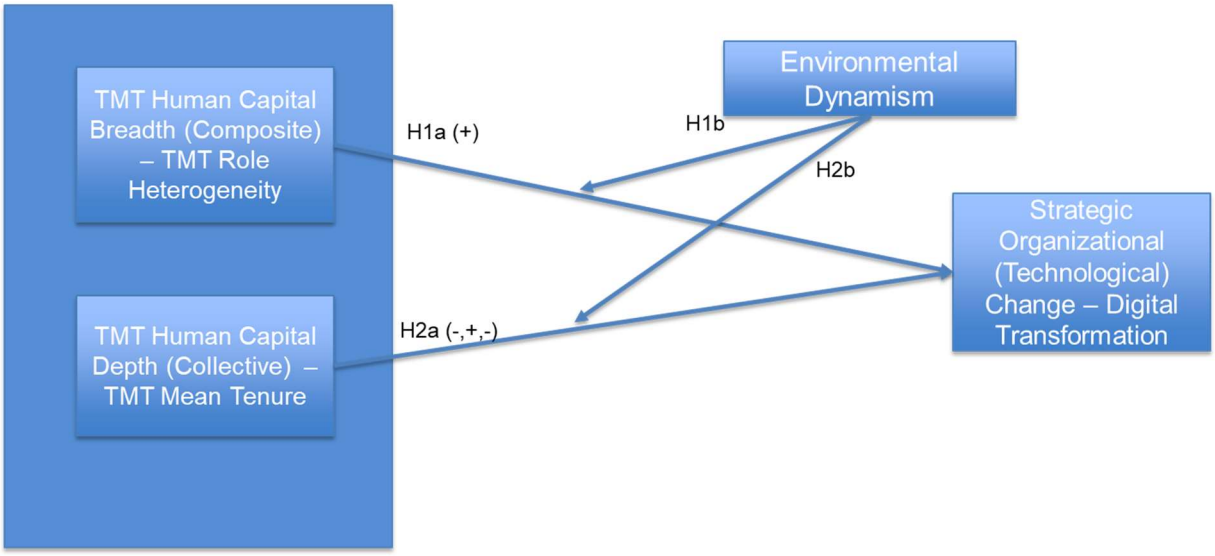
and strategic change initiatives (Finkelstein & Hambrick, 1990). While this may be a general trend, dynamism will affect this relationship. In highly dynamic environments, people having spent longer time are familiar with the nature and characteristics of dynamism and hence more open to experimentation in terms of innovations, newer concepts, and technologies. Using a dynamic capabilities argument, Jiao et al (Jiao et al., 2013) have found that capabilities of opportunity-sensing and reconfiguration facilitate better implementation of strategic organizational change in high dynamism contexts. But people having spent lesser time in the organization need to familiarize themselves with the finer nuances of the organization, hence less open to change (proxied by digital transformation). In less dynamic contexts, however, people with longer tenures would have become too comfortable with status quo, so for them to perceive newer changes will be difficult. Those with shorter tenures have seen the external environment less subjectively as compared to those who have spent longer times in the same organization. So those with shorter tenures are more likely to be open to change.

At lower mean tenures in a less dynamic environment, the negative relationship may get reduced because there are not too many unknowns to be addressed so the decision-making may be comparatively straight-forward. At lower mean tenures in highly dynamic environments, top managers may not have the optimal cohesion, so making any consensual choice will be more difficult. At moderate team tenures, managers will be reasonably well-coordinated with each other, will be able to understand each other better. So whether the environment is dynamic or stable, better synchronization will lead to efficient decision-making. At high team tenures, inter-personal conflicts and politics may have a negative effect on making consensual choices. So, the next moderation hypothesis:

Hypothesis 2(b): The relationship between TMT mean tenure and digital transformation is moderated by environmental dynamism such that in a less dynamic context, the curvilinear relationship will be attenuated and in a highly dynamic context, the relationship will be accentuated

All my hypothesized relationships related to TMT human capital are represented in figure 4 below:

Figure 4: Hypothesized relationships for TMT human capital



CEO Characteristics and Digital Transformation:

CEOs are a part of the top management team, yet they have been studied widely for the influence of their individual characteristics. The large literature on CEO characteristics based on upper-echelons theory is classified through three domains (Busenbark, Krause, Boivie, & Graffin, 2016) – (a) the position (e.g. corporate governance mechanisms, CEO strategic

influence, CEO selection); (b) the person (e.g. CEO identity and firm, CEO personality and characteristics, CEO peer and reference groups); and (c) the environment (e.g. external attributions to CEO, assumptions about CEO, and attention to CEO). For the purpose of my dissertation, I examine only the second domain, "the person" (CEO personal characteristics).

As stated earlier, CEOs' influence on organizational outcomes has grown over the last few decades (Quigley & Hambrick, 2015). In fact, several literary references widely share the belief that CEO is the most powerful organizational member in the modern organization (Daily & Johnson, 1997). CEOs' influence on organizational performance is believed to have increased over the past several decades and the percentage of variance explained by CEO characteristics has increased over this time. A six-decade long analysis conducted by Quigley and Hambrick (Quigley & Hambrick, 2015) found that the mean effect of CEOs on RoS, RoA and MTB (Return on Sales, Return on Assets and Market to Book Ratio) was about 10-12% during late 1960s to early 1980s. This increased to 15-17% during mid 1980s and late 1990s after which the influence has grown to as much as 18-20% until late 2000s.

Certain CEO characteristics like stock ownership, duality (CEOs who also hold board positions), elite education, and functional expertise determine the extent of power they possess. These characteristics have been consolidated under a larger construct called CEO power, has been studied as a composite of 4 reflective constructs – structural power, ownership power, expert power, and prestige power (Daily & Johnson, 1997; Finkelstein, 1992). Structural power comprises of CEO characteristics like CEO duality, board independence, and compensation ratio. Ownership power is measured using percentage of CEO shareholding and founder status. Expert power is measured as number of functional background areas . Prestige power is measured in two ways - number of boards served on, and prestigious educational background. CEOs' sense

of identity coupled with these characteristics in turn determine the extent of power they can have on organizational choices and actions, as has been discussed in the literature background section. If a CEO is powerful, his/ her opinion in decision and execution of strategic choices will be high. This leads to the next hypothesis.

Hypothesis 3: CEO Power is positively associated with digital transformation

Firms most likely to undergo strategic change are characterized by shorter organizational tenures (Wiersema & Bantel, 1992). Past research on CEOs suggests that total organizational tenure influences CEOs' cognitive orientations and knowledge base (Rajagopalan & Datta, 1996). Longer tenures may impede information processing (Miller, D., 1991) and increase cognitive rigidity (Bantel & Jackson, 1989). So, longer a CEO has spent time in an organization before occupying the position (inside CEO), limited information processing and high cognitive rigidity will in turn impede the inducement to bring strategic change. Explained differently, higher firm tenures have been found associated with firms' tendencies to persist with past strategies, thus resisting strategic change (Finkelstein & Hambrick, 1990; Hambrick, Geletkanycz, & Fredrickson, 1993). Past literature also suggests that top executives' longer tenures in their organizations are associated with developing high degree of organization-specific knowledge (Gupta, 1984). This high level of organizational knowledge comprises of everything that the firm may have faced in the past, including the industry characteristics of the industry to which the firm belongs. If the industry is stable, firms are more likely to persist with past strategies, because changes in strategy in a stable context reduces efficiency of the strategy

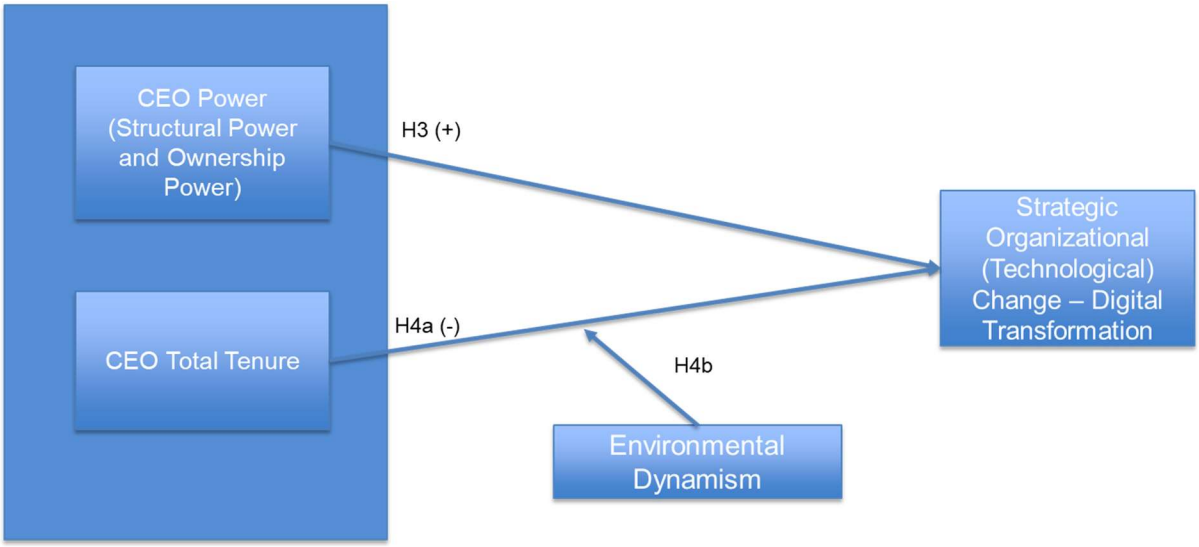
followed. This is substantiated by past research, which suggests that industry stability is related to efficiency-oriented strategies (Chaganti, Rajeswararao & Sambharya, 1987; Thomas, Litschert, & Ramaswamy, 1991). So, in stable or low dynamism industries, longer tenured CEOs are expected to pursue continuity and resist change. Thus, there will be a stronger negative relationship between CEO tenure and strategic change. But in high dynamism industries, relationship of longer tenured CEOs is expected to manifest differently. In high dynamism industries, firms need to continuously monitor the industrial environment and keep modifying the strategic response to cope up, as suggested by the dynamic capabilities perspective. Therefore, even though longer tenured, CEOs will be more open to making changes that will keep the firm abreast of environmental idiosyncrasies. Therefore, the magnitude of negative relationship of tenure to digital transformation is expected to reduce.

Hypothesis 4(a): CEO Organizational tenure is negatively associated with digital transformation

Hypothesis 4(b): Relationship between CEO organizational tenure and digital transformation is moderated by environmental dynamism such that lower dynamism accentuates the relationship whereas higher dynamism attenuates the negative relationship

My hypothesized sub-model for CEO characteristics is given in figure 5 below:

Figure 5: CEO characteristics and digital transformation


CEO Power (Structural Power and Ownership Power)
CEO Total Tenure
H3 (+)
H4a (-)
H4b
Environmental Dynamism
Strategic Organizational (Technological) Change – Digital Transformation

CHAPTER 4: METHODS

To study how companies are responding to the digital transformation initiatives, I analyzed a panel of the S&P 500 firms from 2008 to 2017. This period corresponds to the rise of digital transformation in U.S. industry. In the early 2000s Google's search engine rose to prominence along with Amazon Web Services, which provided data on website popularity, Internet traffic patterns and other statistics for marketers and developers (Miller, F., Vandome, & McBrewster, 2010).. The concept of grid computing, first introduced in 1990s, transformed into SaaS (software as a service) by 2000 and further into cloud computing in 2007 (Böhm, Leimeister, Riedl, & Krcmar, 2010). The term business intelligence and analytics, which subsequently became to be known as big data analytics, first became popular after the popular best-seller "Competing on Analytics" (Davenport & Harris, 2007) was published.

For all the basic financial data and some demographic data I used Compustat. For additional data about TMT members and their characteristics, I used the BoardEx database. My interest was to study top management team members' characteristics. In several companies, some top management team members were also board members. So to isolate the effects of board characteristics, I decided to use only those top executives' data who were not board members. However, for CEOs I took all CEO data, irrespective of whether they were board members or not. For my measures related to digital transformation, I used company annual reports/10K filings extracted from the website annualreports.com. Some of the missing reports were extracted manually from the websites of the respective companies.

MEASURES

Dependent Variable

Digital Transformation (DT). Digital transformation as a construct for academic study is a recent concept and without much empirical literature. To develop a measure I used a text-analytic approach to measure the extent of digital transformation in companies by the degree to which the technologies and processes used in digital transformation appeared in the annual reports (Daft & Lengel, 1986; Orlikowski, 1993). This was operationalized by measuring usage of related terms or synonyms in 10K filings and annual reports (Bowman, 1984; Cho & Hambrick, 2006; Clapham & Schwenk, 1991). Text appearances of words and phrases in annual reports has been previously used to measure constructs including corporate strategy (Bowman, 1984), causal attributions (Clapham & Schwenk, 1991), and attention focus (Nadkarni & Barr, 2008) among others.

This technique is an appropriate proxy for digital transformation for several reasons. First, the appearance of certain words and phrases in documents such as the annual report signals awareness of digital transformation by senior management. In fact, "statements by management, in annual reports and elsewhere, provide some of the best data on the cognitive aspects of strategic management" (Clapham & Schwenk, 1991, p 219). Clapham et al (1991) argue that annual reports indeed act as references towards understanding how external events and executives' own actions may impact their firms' performance (Bowman, 1976; Bowman, 1978; Bowman, 1984; Salancik & Meindl, 1984; Staw, McKechnie, & Puffer, 1983). I argue that if a company is including terms associated with digital transformation it means that top managers have been thinking about this phenomenon and executives are cognitively aware of digital

transformation. This is a content element of change in the multi-lens model of strategic change (Rajagopalan & Spreitzer, 1997).

Secondly, higher frequency of words and phrases related to digital transformation suggests that it is higher in organizational priorities for a company relative to others with fewer mentions. “Words that are frequently used are cognitively central and reflect what is most on the user's mind; words that are used infrequently or not at all are at the cognitive periphery, perhaps even representing uncomfortable or alien concepts (Huff, 1990)” (Cho & Hambrick, 2006; p 459). If organizations are writing about digital transformation more frequently in their annual reports it suggests that such organizations have digital transformation among their top priorities. To explain this in the context of cognitive model of organizational renewal (Barr et al, 1992), cognitive or mental models of managers must change in order to cope up with environmental changes. The frequency of the usage the terms or increase in the frequency in usage may suggest that managers are paying attention to and making sense of or interpreting that digital transformation is something that they must think about (Gioia & Chittipedi, 1991), and then changing their mental maps to align their thought process about external environment, which in turn may lead to some discussion and decision leading to digital transformation initiatives.

Thirdly, going beyond the cognitive model of organizational renewal, past research using causal mapping techniques have found a strong correlation between managerial cognition as reflected from letters to shareholders and annual reports, and strategic responses to environmental changes (Nadkarni & Barr, 2008). This means that the content of annual reports may strongly reflect the actions taken by top management.

Method for extracting text data and creating the variable. The creation of a measurable variable for digital transformation was carried out in 3 phases: (a) Creation of a custom dictionary; (b) creation of the data corpus comprising of the annual reports of the relevant companies; (c) analyzing the annual reports by parsing words from the custom vocabulary using text analytics processes.

(a) *Custom dictionary*. Before measuring the occurrence of terms associated with digital transformation, I created a custom dictionary for defining the domain of digital transformation using three different ways or sources. The first was a list of key-words from various popular practitioner literature sources, as identified using a google search. The second way was to perform an automated textual analysis of academic literature related to digital transformation. The third way was to perform a content analysis of few popular books on digital transformation.

The first method was to perform a google search for "glossary of terms related to digital transformation". Using google search, I identified top 14 different websites which had a list of words associated with digital transformation. This list of websites and some sample terms extracted from those websites is given in Appendix A.

INSERT APPENDIX A ABOUT HERE

I manually compiled all the key-terms given in these different websites and recorded the frequency of the terms that featured in each of these websites. This corpus of 229 phrases became my primary data source. The second method was to extract such key-terms from

academic literature on digital transformation. This was done using the Web of Science database (owned by Clarivate Analytics) on Sept 12th, 2019. Using "digital transformation" as the search term, I downloaded abstracts of all the 690 articles and conference papers that were identified. Key information about few of those articles is given in Appendix B1 and some sample terms in Appendix B2.

INSERT APPENDIXES B1 AND B2 ABOUT HERE

This corpus of articles was then subjected to a text analytic process to give out list of key-words based on their frequency of occurrence in those articles. After cleaning out some noisy data, I got a list of 294 phrases which occurred at least 3 times or more in the academic literature. The third method was to identify top rated books related to digital transformation. I used a website called goodreads.com to identify the top 10 highest rated books on digital transformation, out of which I shortlisted four books whose pdf files were publicly available (Brynjolfsson & McAfee, 2014; Greenway, Terrett, Bracken, & Loosemore, 2018; Rogers, 2016; Westerman, Bonnet, & McAfee, 2014a). I also downloaded three other books whose copies were publicly available in pdf form (Andersson, Movin, Mähring, Teigland, & Wennberg, 2018; Kerravala & Miller, 2017; Palmer et al., 2019). I extracted about 150 top frequently occurring terms obtained from this analysis. A list of these books and their sources is given in Appendix C1 and some sample terms from this method is given in Appendix C2.

INSERT APPENDIXES C1 AND C2 ABOUT HERE

A combination of these three sources gave a list of about 400 key-terms or phrases. For validation of the list and trimming it to make them more relevant, I used the SME determined keyword approach - creating a glossary of key terms based on subject matter experts' opinion on a list of pre-determined terms and search public data of these companies for frequency of usage. To enable seek opinions from subject matter experts, I procured consent from the IRB Review Board. After receiving their consent, I sent the list of keywords to about 25 practitioners within the domain of information technology who work on products and solutions related to digital transformation. The feedback sheet contained the key-terms as well as columns for rating those terms on a scale of 0 to 5. In the sheet I asked the question : "On a scale of 0 to 5, please rate the following terms on their association with Digital Transformation , where 0 = Not Associated and 5 = Highly Associated". I also added the comment: "In case of any ambiguity, you can rate the term/ phrase based on whatever context comes to your mind after you see the word. If you feel some word/s is/are left out, you may add them into columns C, D, and E on row 116. If you want to add more than 3 words, please add them in rows 117 onwards". In addition, I also asked 3 questions to get some demographic information about the respondents : (1) Total years of experience; (2) Academic Scholar OR Corporate Professional OR Both; (3) Industry/ Industrial Sector currently associated with; (4) Your Role OR Type of digital transformation-related activity you are associated with (in a single sentence). A sample template sent to these subject matter experts is given in Appendix D.

INSERT APPENDIX D ABOUT HERE

I got responses from 8 experts. After receiving these responses, I purged the list using my judgement about the average lower-scored keywords. Generally, all the words that were rated at an average of less than 1.75 were removed from the analysis. This purged list was then sent across for another additional round of scrutiny to two scholars who work in the domain of digital transformation among other topics. With this additional scrutiny, the final list of keywords comprised of 348 terms, down from 461 words in the original list. This final set of keywords is featured in Appendix E.

INSERT APPENDIX E ABOUT HERE

(b) Creation of data corpus - Downloading annual reports. As mentioned earlier, I used annual reports of companies to identify keywords based on which my dependent variable was be measured. I downloaded most of my data from the website annualreports.com. As this was done using an automated web-scraping system with the help of a graduate student, I downloaded all the data that were available on this website. The Python code used for extracting these annual reports is given in Appendix F.

INSERT APPENDIX F ABOUT HERE

The full data comprised of annual reports from my group of interest (S & P 500) as well as other listed companies. From the full dataset, I extracted the annual reports of my interest. Out of the total of 740 companies that were included in the S & P 500 index during the period 2008 to 2017, I was able to compile a database comprising of 390 companies. There were several reasons for this shrinkage. First, I wanted only those companies which were included in the index for at least 7 years during the period of interest. Second, some companies from withing such 7-year category were either acquired by or merged with some other company. Third, for several companies, annual reports were not available at all on the annualreports.com website. From among the remaining 390 companies, some years' annual reports were not available on the annualreports.com website. Such were downloaded from the websites of the respective companies with the help of services of another graduate assistant from the Management department.

(c) Content analysis using text analytic approach. With the help of a graduate student from the Information Systems department, the annual reports were analyzed using a program in Python language for the occurrence of these keywords. The program compared all the annual report data against the created and validated custom dictionary and scored each of the terms occurring in the annual reports against the dictionary availability based on frequency of occurrence. This program could read and score annual reports of 374 companies for all the available documents for the period 2008 to 2017. The Python code for this compilation is given in Appendix G.

INSERT APPENDIX G ABOUT HERE

The output consisted of the score of occurrence of each term in each document, the total count of such occurrences per document (ranging from 0 to 10593), and the total number of words in each document (ranging from 1 to 557134). Upon verifying with some of the low total word count documents, I found that some of these files could not be read fully by the program whereas some others had just 1-2 pages of the report. As the variance was huge with such unusual counts, I cleaned the data by eliminating all the documents which had less than 100 total words. There were few other companies for which annual reports for all 10 years were not available. So data for a total of 3709 files were finally used in my analysis. An extract of company-year-wise summary of the number of terms and their aggregate is given in Table 1.

INSERT TABLE 1 ABOUT HERE

For the purpose of standardization, I took a ratio of the total occurrences to the total word count per document. This gave me a wide range of ratios ranging from 0 (documents in which no dictionary term occurred) to 0.033. These ratios too had a wide negative skew towards 0, so I transformed the values by multiplying them by 10000 and then taking log of those values. This transformation yielded a normalized set of observations.

Independent Variables

CEO characteristics and TMT characteristics were collected separately primarily through the BoardEx database. Non-CEO top management team members were included in calculating TMT Role Heterogeneity and TMT Tenure . I could extract 435540 observations belonging to 40089 top management executives spread across 571 companies over a 10-year period from

2008 to 2017. Many of these company data were dropped to match with the dataset of my dependent variable (390 companies' annual reports). Following is my list of variables and their operationalizations.

TMT Role Heterogeneity. Role heterogeneity is defined as the total number of unique top management jobs held by members of the top management team divided by the total number of members of the top management team in a given year as reported in BoardEx. For TMT role heterogeneity, I had 3233 observations ranging from 1 to 386 with a mean of 55.8. I eventually log-transformed this variable to eliminate skewness by using the zero-skewness log transform command in Stata16.

TMT Tenure. This was calculated as mean of firm tenure of all the TMT members. The individual tenure of each top management team member was measured in number of days and the mean was taken for all members of the TMT. The 3224 observations ranged from 17 to 4212 with a mean of 1338. To eliminate this wide skew, I transformed this variable using the zero-skewness log transform command in Stata16.

CEO Power. This was measured using Daily & Johnson's (Daily & Johnson, 1997) approach of structural power and ownership power. All these measures were created from the Execucomp-Compustat database available through WRDS. Following are operationalizations of CEO power.

CEO Duality (Structural Power). CEO duality is measured as a dichotomous identifying and tagging all the CEOs listed in ExecuComp that had Chairperson titles. CEO titles which had variations of the word-stub "Chair", i.e. "Chairman" or "Chair" or "Chairperson" in the title

were coded as 1. Out of 3272 observations, 1398 CEOs entries were not holding dual responsibility whereas 1874 entries had board duality.

Percentage of shares owned by CEO (ownership power). This is defined as the percentage of the CEO's shares to total shares outstanding and was drawn from Execucomp-Compustat. As per the data of 2975 observations used in my analysis, the values ranged from 0% to 53.83% with a mean of 1.22%. As there was high skew in the observations, I transformed this variable by first multiplying it by 100 and then taking a natural log of the new values.

Founder status (ownership power). Founder status is defined as whether the CEO is a founder or not of firm drawn from Execucomp. I created this binary variable by identifying and tagging all the mentions of the variants of the term "co-founder" (co founder, cofounder) in the titles of the respective CEOs, marking a 1 for presence and 0 for absence of the term. I got 103 entries for co-founder status in my dataset.

CEO Total Tenure. This is the total duration the CEO has worked in the focal firm (including pre-CEO tenure)drawn from the Execucomp-Compustat database. This was calculated as number of days in the position. Due to some inconsistencies (minimum tenure was negative), I had to winsorize the data. Against the pre-normalized range of -418 to 13879, the new range I obtained was 200 to 11027 for a total of 3241 observations. As the data-points were skewed, I transformed these into log using the zero-skewness log transform function in Stata16.

Control Variables Definitions

My control variables were firm age and firm revenue. Firm age may influence the results in several ways. First, an older firm may belong to an industry which may have been in existence for a long time. It may be possible that such older industry is less dynamic due to its very nature

(for instance mining). Alternately, a newer firm which was founded just a few years back may belong to an industry which by is very nature may be highly dynamic (for instance IT services). Second, an older firm may have long serving top management team members. For instance, a company that has been in existence for more than 50 years may have people working with such firms for 20 years or 25 years or 30 years. In contrast, a newly established firm with just 10 years of history may have top managers with highest tenure of 10 years. This too may confound the actual influence of my independent variables.

Firm revenue is another control variable I used. One the one hand, firm revenue represents size of the firm. On the other hand, it represents the financial and other resources at its disposal. Both these factors may influence the decisions of top executives related to strategic change or digital transformation. Bigger size may automatically prompt firms to consider such measures that help them manage their businesses more efficiently. Alternately, bigger size may also interfere with quick decision-making due to the inherent structural inertia. Both these effects can be controlled with the approach of controlling for firm revenue. Another reason is financial resources. Higher revenues may represent more financial resources. Due to this, top executives may find it easier to take decisions related to capital-intensive investments, like introduction of ERP systems or automation. Controlling for revenues can thus help control the effect of higher absolute value of resource allocation. Both, firm age, and firm revenue related data were extracted from CRSP-Compustat.

Firm age. This is calculated as the difference between the year of founding and 2017. Out of 3887 observations, my data ranged from 3 years to 233 years with a mean of 69 years.

Firm revenues. Total revenue for the company annually as recorded in CRSP-Compustat. The revenue ranged from USD 3.92 million to 483,521 million with a mean of 17518 million. This too was log-transformed to normalize the data.

Moderating variable

***Environmental Dynamism*.** This variable was operationalized using Dess & Beard's (Dess & Beard, 1984) industry-based dynamism measure, defined as volatility of the rate of change of annual industry sales. This was calculated by dividing the standard error of the rate of change of annual sales by the mean annual sales (Barry et al., 2006) for each 3-digit NAICS industry. To create the basis for calculation of dynamism at the industry level, I extracted key firm data from the CRSP-Compustat database. This dataset comprised of 54930 observations pertaining to 8438 firms through the years 2008 to 2017 firms belonging to 146 different industries spread over a range from 2 digit to 6 digits according to the NAICS classification. . My dataset comprised of 370 companies. So, based on an optimal match of master data versus sample data, I created my classification of industry based on 3-digit NAICS codes which finally yielded 7360 companies corresponding to the industries in which my sample companies were nested in. This enabled me to have a range of 5 companies per industry to 993 companies per industry in the master, as compared to range of 1 company per industry to 46 companies per industry in my sample data, i.e. the ratio ranged from 1:5 to 21:122. These ratio ranges were sufficient for calculating environmental dynamism indexes per industry. Using the master dataset that comprised of 8348 firms distributed amongst 146 industries with data for the period 2008 to 2017, I proceeded to calculate the environmental dynamism index in the manner as explained previously. This industry index was then assigned to each company within the industry, i.e., all the companies in my sample data set that belonged to the same industry were assigned an

environmental dynamism index that remained the same for each company during a particular year. The final mean values for industries over the 10-year period ranged from 0.0128 for NAICS code 531 (Real Estate) to 0.0989 for NAICS code 434 (Non-store retailers). A list of industries is given in Table 2.

INSERT TABLE 2 ABOUT HERE

Descriptive Statistics and Correlations for all variables

A summarized chart of descriptive statistics of the raw values of all variables is given in table 3. As mentioned in the individual variable operationalizations earlier, these were subsequently transformed for normalization and better interpretation.

INSERT TABLE 3 ABOUT HERE

For initial assessment of the relationships between different variables, I carried out a correlation analysis for them. This is presented in table 4. All the variables of interest have significant correlations ($p<0.05$) with the log transformed word count variable (the dependent variable which represents digital transformation). These correlations range from a lowest of 0.042 for log transformed CEO shareholding × 100 variable to a highest of 0.205 for log transformed TMT role heterogeneity.

INSERT TABLE 4 ABOUT HERE

Analytical Approach

The final panel dataset for analysis included 370 companies representing 65 industries over a period from 2008 to 2017. The data are structured hierarchically such that the yearly observations for each company are nested within the company and multiple companies are nested within industries (Arceneaux & Nickerson, 2009; Steenbergen & Jones, 2002). Therefore, I tested my hypotheses using hierarchical linear modeling (HLM)/ multi-level linear modeling (MLM).

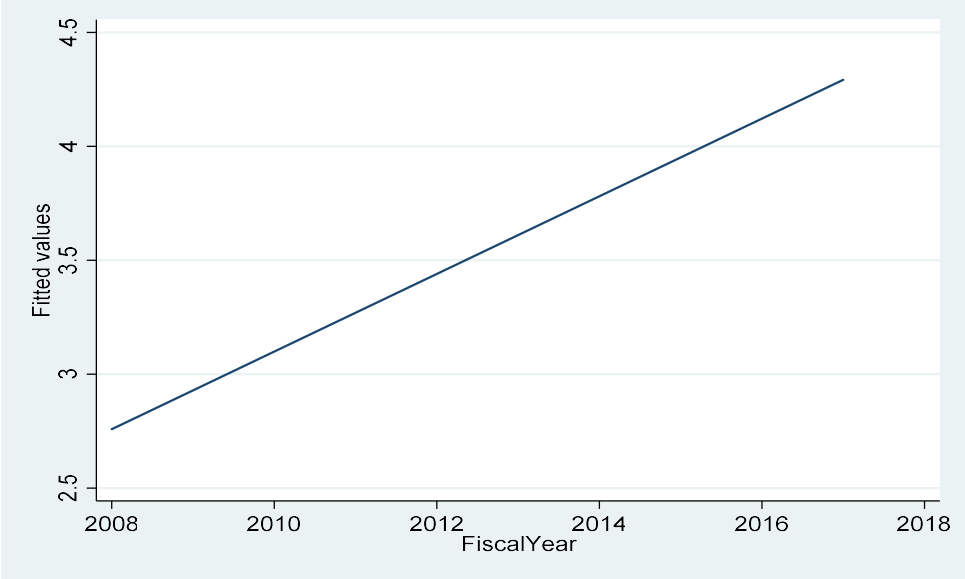
Another reason for choosing HLM is that it facilitates identification and isolation of group heterogeneities from individual heterogeneities (Nielsen & Nielsen, 2013). The HLM model permits the comparison of company effects on digital transformation and industry effect on digital transformation, i.e. by specifying separate random components for different levels of analysis (Hair Jr & Fávero, 2019; Heck, Thomas, & Tabata, 2009). I describe my analysis in the next chapter.

CHAPTER 5 – ANALYSIS, RESULTS, AND INTERPRETATION

As mentioned earlier, my final dataset comprised of an unbalanced panel of 370 companies spread through 65 different industries (classified on the basis of NAICS 3-digit codes) with number of time observations ranging from 5 years to 10 years and number of companies per industry ranging from 1 to 46.

I used the MLM (multi-level modeling) feature in Stata to run my analysis. One test to identify whether MLM is the appropriate analytical approach is by graphing the year trends against digital transformation scores between different companies. Figure 6 shows trend lines of a few sample companies from within the dataset. The graph shows that the temporal evolution of digital transformation present different intercepts and slopes between companies.

Figure 6: Temporal Evolution of Digital Transformation In Sample

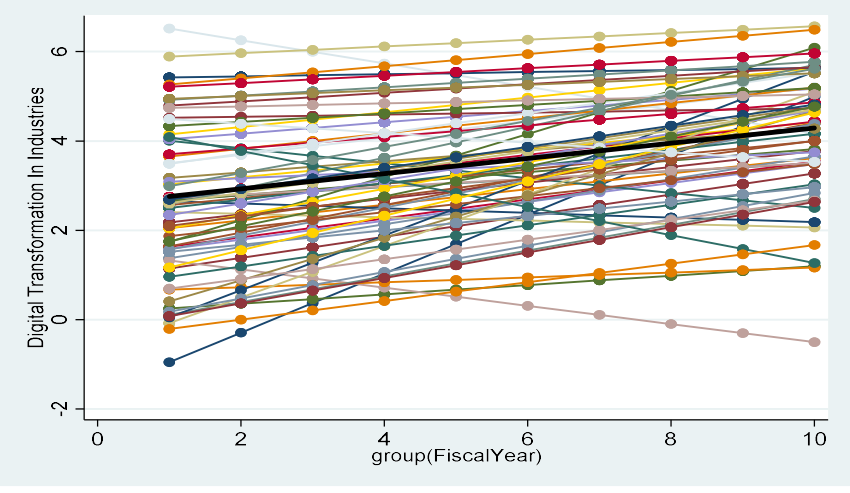


These different intercepts and slopes suggest that a 2nd hierarchical level comprising of company-wise intercept and slope random effects is justified, thus constructing the 2-level model.

My theoretical model also incorporates the industry effects of environmental dynamism on the digital transformation efforts of companies. To empirically establish the necessity of thus incorporating a third level, another graphical representation can be shown.

Figure 7 below shows the temporal evolutions of the average industry score for digital transformation. The changing trend over time further justifies for applying a three-level model hierarchical model. This figure also demonstrates the linear adjustment of digital transformation score of the industry over time through OLS. Additionally, it shows intercept and slope random effects that will be estimated at the third level, i.e. industry level, because each industry presents different intercepts and slopes.

Figure 7: Temporal Evolution of Average Digital Transformation Scores of Companies In Each Industry



With the theoretical and empirical justification that establishes that the data structure is characterized by temporal nesting of companies from different industries with different values of digital transformation scores over time, I proceeded to estimate the null model. With this estimation, I wanted to determine if digital transformation scores varied between companies in the same industry and companies in different industries. This model is called null model because it does not include any explanatory variables. It considers only one intercept with 3 different error terms u_{00k}, r_{0jk} and e_{tjk}, having variances, respectively, equal to Tau_{u000}, Tau_{r000} and $Sigma^2$.

This model has the expression as given below.

Null Model with Level 2 and Level 3 Random Intercepts or RIs (Model 0)

$(\text{Digital Transformation Score})_{tjk} = Pi_{0jk} + e_{tjk}$; where

$Pi_{0jk} = b_{00k} + r_{0jk}$; where

$b_{00k} = Gamma_{000} + u_{00k}$

So the full null model is : $(\text{Digital Transformation Score})_{tjk} = Gamma_{000} + u_{00k} + r_{0jk} + e_{tjk}$

Here, the character Pi_{0jk} represents the random intercept for company j in industry k, i.e. the group mean of digital transformation score for company j in industry k. The error term e represents the error associated with a specific observation of digital transformation score for this company j at time t. Because the company is nested within industry k, the random intercept Pi_{0jk} may be decomposed into a random intercept b_{00k} representing the group mean of industry k and an error term r_{0jk} representing the error value for the specific observation of company j in industry k. Further, as my dataset comprises of multiple industries, the random intercept b_{00k} too

can be decomposed further into two parts - a constant $Gamma_{000}$ that represents the overall grand mean of all digital transformation scores, and an error term u_{00k} to represent a specific value for industry k. Hence the full null model.

Upon estimating the null model (figure 8 below), I got a value of 3.39 (s.e. 0.157) for the parameter $Gamma_{000}$, that forms the general intercept representing the average of the overall expected digital transformation score, a fixed effects component.

Figure 8: Null Model with Level 2 and Level 3 Random Intercepts (RIs)

Mixed-effects REML regression — Number of obs = 3,643

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
NAICS_Code~t	65	9	56.0	419
GVKey	370	5	9.8	10

Log restricted-likelihood = -6317.6721 — Wald chi2(0) = . — Prob > chi2 = .

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	3.389929	.1568711	21.61	0.000	3.082467	3.69739

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Identity var(_cons)	1.135664	.2842215	.6953776	1.854724
GVKey: Identity var(_cons)	.9449735	.0888628	.7859137	1.136225
var(Residual)	1.49596	.0369793	1.42521	1.570223

LR test vs. linear model: chi2(2) = 2155.21 — Prob > chi2 = 0.0000

So the estimated null model is : $(\text{Digital Transformation Score})_{tjk} = 3.39 + u_{00k} + r_{0jk} + e_{tjk}$

The variances of the error terms were estimated at 1.136 (Tau_{u000}) for industry, 0.945 (Tau_{r000}) for company, and 1.496 ($Sigma^2$) for company residual. With the 2 different nested levels, 2 intra-class correlations (ICC) were identified. The first one was the correlation between the digital transformation scores at different time periods for company j from an industry k ($Rho_{company|industry}$) with a value of 0.581. The second one was the correlation between data of digital transformation scores of different companies j and $j^{/}$ from a certain industry k, having a value of 0.318. This estimation is shown in figure 9 below:

Figure 9: Intra-class correlation between level 2 and level 3 groups in null model

Intraclass correlation

Level	ICC	Std. Err.	[95% Conf.	Interval]
NAICS_Code_3Dgt	.3175264	.0558774	.2191642	.4354174
GVKey|NAICS_Code_3Dgt	.5817366	.0341368	.5137292	.6467725

This may be interpreted such that 31.8% of the random variance in digital transformation score of a company is explained by industry effects and a total of 58.1% of the total random variance can be explained by industry and company effect. As the estimate values are all statistically significant, it indicates that there is significant variation in digital transformation scores between companies and between industries. This provides additional basis for choosing multi-level modeling instead of OLS regression. Estimating the null model also provided me with the additional information about the variance that occurred in digital transformation scores between companies of the same industry and between companies from different industries over time during the full period of study.

My hypotheses relate to examining company specific characteristics that influence changes in digital transformation scores between companies over time and the moderating effect

of environmental dynamism. The null model or Level 0 was the first step towards that objective. The next step in my analysis was to investigate the effects of time on digital transformation scores of companies, and whether those effects followed a linear pattern. This may be called the linear trend model with random intercepts.

Linear Trend Model with Random Intercept (RI): (Model 1 with Fixed Intercept or FI, Time Slope or TS, and Level 2 & 3 RIs)

$$(\text{Digital Transformation Score})_{tjk} = \text{Pi}_{0jk} + \text{Pi}_{1jk.} \cdot \text{FiscalYear}_{jk} + e_{tjk} \text{ ; where}$$

$$\text{Pi}_{0jk} = b_{00k} + r_{0jk}\text{; and}$$

$$\text{Pi}_{1jk} = b_{10k}\text{; where}$$

$$b_{00k} = \text{Gamma}_{000} + u_{00k}\text{; and}$$

$$b_{10k} = \text{Gamma}_{100}$$

So the random intercept model is :

$$(\text{Digital Transformation Score})_{tjk} = \text{Gamma}_{000} + \text{Gamma}_{100} \cdot \text{FiscalYear}_{jk} + u_{00k} + r_{0jk} + e_{tjk}$$

Where, Gamma_{000} represents the fixed intercept made by the overall grand mean value of digital transformation scores, Gamma_{100} represents the fixed slope for the grand mean, u_{00k} represents the random intercept for industry k, r_{0jk} represents the random intercept for company j in industry k, and e_{tjk} random error representing digital transformation score for a given company j in industry k at time t.

The estimate for parameter $Gamma_{100}$ (grand mean of digital transformation score for a specific year) is 1.686 and statistically significant, which indicates that a company's digital transformation score follows a linear trend (as per figure 10 below). With this parameter, my random intercept model had the following specification:

$$(\text{Digital Transformation Score})_{tjk} = 2.461 + 0.169.\text{FiscalYear}_{jk} + u_{00k} + r_{0jk} + e_{tjk}$$

The random intercepts for industries and companies are graphically visualized in Figure 11.

The estimates for variances of the error terms were estimated at 1.134 (Tau_{u000}) for random intercept of industry, 0.968 (Tau_{r000}) for that of company, and 1.237 ($Sigma^2$) for company residual. The ICC for level 3 (industry) was 0.339 and that of level 2 (company|industry) was 0.629. So, with the inclusion of a fixed time effect, these ICCs were higher than those of the null model (0.317 and 0.581 respectively – figure 11 versus 12). The variance values also indicate that there is significant variance of intercepts between companies within same industry and companies from different industries.

Figure 10: Linear Trend Model with Random Intercept (RI): (Model 1 with Fixed Intercept FI, Time Slope TS, and Level 2 & 3 RIs)

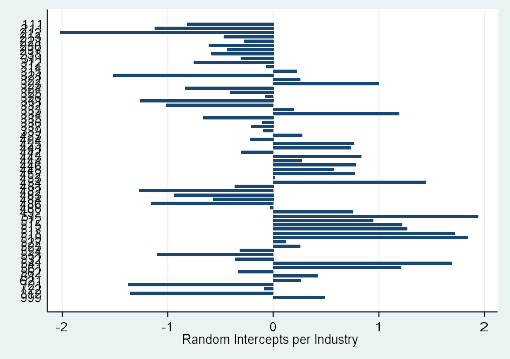
Wald chi2(1) = 686.12
Log restricted-likelihood = -6010.1862 Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1686391	.0064381	26.19	0.000	.1560206	.1812576
_cons	2.461639	.160696	15.32	0.000	2.146681	2.776598

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Identity				
var(_cons)	1.13413	.2839833	.6942631	1.852685
GVKey: Identity				
var(_cons)	.9684799	.0886222	.8094688	1.158727
var(Residual)	1.237304	.0305902	1.178778	1.298736

LR test vs. linear model: chi2(2) = 2511.69 Prob > chi2 = 0.0000

Figure 11: Random intercepts for industries and companies



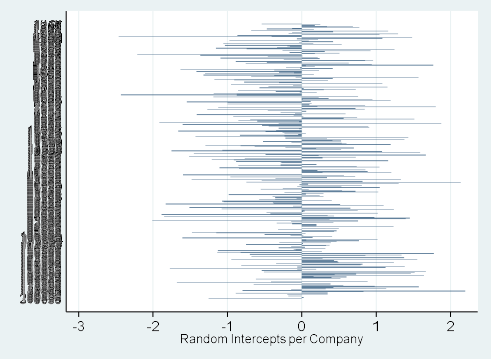


Figure 12: Intraclass correlation for linear trend model with random intercepts only

Residual intraclass correlation

Level	ICC	Std. Err.	[95% Conf.	Interval]
NAICS_Code_3Dgt	.3395686	.0580774	.2363507	.4606708
GVKey|NAICS_Code_3Dgt	.6295401	.0323422	.5642533	.6904122

That gives rise to the next step in my analysis. While different companies have different intercepts or group means for digital transformation scores within an industry, I needed to verify whether there was significant variance of slopes throughout time between these companies. In other words, how much did digital transformation score trends vary over time amongst companies within industries? Hence I decided to add slope random effects to my model, thus making it a linear trend model with random intercepts and slopes.

Linear trend model with random intercepts (RI) and random slopes (RS) (Model 2 with FI, TS, Level 2 RI & RS, Level 3 RI)

To build the next level of my model, in addition to the fixed effects component (Gamma), a variable Mean-Centered Environmental Dynamism was applied to random effects components of company. This was to facilitate separation of the unique effect of industry level environmental dynamism on firm random effects. Figure 13 shows the results estimated from this model. There is a very miniscule change in the fixed effects estimates ($Gamma_{000} = 2.44$; $Gamma_{100} = 0.162$); but the variance estimates are different from the previous model. A Chi^2 test for model comparison between the null model (figure 5 random intercepts) and the current model (figure 8) using the formula ($-2 \cdot LL_{r\text{-randomintercept}} - (-2 \cdot LL_{r\text{-randomslope}})$) gives a value of 735.688, significant at

the 0.05 level. This test indicates that a linear trend model with and random intercepts and slopes is an appropriate model.

Figure 13: Linear trend model with random intercepts (RI) and random slopes (RS)

Mixed-effects REML regression — Number of obs = 3,554

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
NAICS_Code~t	65	7	54.7	410
GVKey	369	1	9.6	10

Log restricted-likelihood = -5805.17 — Wald chi2(1) = 551.99; Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1732297	.0073732	23.49	0.000	.1587785	.1876809
_cons	2.444327	.1624803	15.04	0.000	2.125872	2.762783

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Identity				
var(_cons)	1.139921	.2863917	.696659	1.865216
GVKey: Independent				
var(Cen_En~s)	5421.818	1167.737	3554.802	8269.409
var(_cons)	.9935132	.0898077	.8322046	1.186089
var(Residual)	1.082789	.0298264	1.02588	1.142854

LR test vs. linear model: chi2(3) = 2572.91 — Prob > chi2 = 0.0000

This analytical model reflects my theoretical model that companies' digital transformation follow linear trends throughout time (fixed effects are significant). Additionally, it also reflects that there are random variances among companies within and between industries (random effects parameters are significant).

Linear Trend Model with Random Intercepts and Slopes and with Control Variables or CVs (Model 3 with CVs, TS, Level 2 RI and RS, Level 3 RI)

Now I specify my model with my control variables at the first level of the model. Age of the firm may determine how it conducts business, what kind of people work in those firms, which type of business it is into, among others. Besides, it has been used as a control variable in almost all firm level research, hence I decided to include it. Another very frequently used control mechanism is that of firm size. Among the different ways to operationalize, firm sales has been very frequently used (mostly in a log-transformed form), hence I decided to use the same operationalization. As is seen in Figure 14, there is a marginal increase in fit as compared to the previous model.

Figure 14: Linear Trend Model with RIs and RS and with Control Variables

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
NAICS_Code~t	65	7	54.7	410
GVKey	369	1	9.6	10

Log restricted-likelihood = -5804.816

Wald chi2(3) = 570.67
Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1652997	.007669	21.55	0.000	.1502687	.1803307
Firm_Age	-.0031238	.0013011	-2.40	0.016	-.0056738	-.0005737
ln_FirmSales_CRSP_Cmpstat	.1528655	.0415849	3.68	0.000	.0713606	.2343704
_cons	1.326117	.3939826	3.37	0.001	.5539256	2.098309

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Identity				
var(_cons)	1.064799	.2686716	.6493684	1.745999
GVKey: Independent				
var(Cen_En~s)	5314.403	1152.541	3474.189	8129.346
var(_cons)	.9695965	.0879606	.8116543	1.158273
var(Residual)	1.082304	.0298115	1.025424	1.142339

LR test vs. linear model: chi2(3) = 2462.29 Prob > chi2 = 0.0000

This model indicates that the fixed intercept has reduced drastically to accommodate the fixed effects of the slope created by the control variables. In this model, the random effects

(intercepts and slopes) are still significant. Therefore in the next stage, I introduced all my explanatory variables amongst which I have hypothesized four direct relationships.

Full Model with all explanatory variables (Model 4 with All Variables, TS, Level 2 RI & RS, Level 3 RI)

A comparison of log-likelihood ratio suggests that this model is a better fit over the previous control variables only model (figure 14 versus figure 15: from -5804 to -4403).

Figure 15: Full Model with all explanatory variables

Mixed-effects ML regression — Number of obs = 2,729

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
NAICS_Code~t	55	5	49.6	329
GVKey	315	2	8.7	10

Log likelihood = -4403.6475 — Wald chi2(10) = 418.74; Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1528674	.0109477	13.96	0.000	.1314104	.1743245
Firm_Age	-.0037484	.0014669	-2.56	0.011	-.0066234	-.0008734
ln_FirmSales_CRSP_Cmpstat	.0606092	.0545144	1.11	0.266	-.046237	.1674554
Ln_Role_Hetero	.2301237	.1008008	2.28	0.022	.0325577	.4276897
Ln_TMT_MeanTenure	.7007257	5.024978	0.14	0.889	-9.14805	10.5495
c.Ln_TMT_MeanTenure#c.Ln_TMT_MeanTenure	-.0354069	.3302851	-0.11	0.915	-.6827538	.61194
ln_0_CEO_TotalTenure_0	-.0727489	.1113403	-0.65	0.514	-.2909718	.145474
CEO_Pwr1_Duality_ExeCompu	-.0569669	.0703558	-0.81	0.418	-.1948617	.0809279
Log_CEO_Pwr2_ShrOwnPrcntx100	.0218378	.0241845	0.90	0.367	-.0255631	.0692386
CEO_Pwr3_FndrStat_ExeCompu	-.387258	.2476009	-1.56	0.118	-.8725469	.0980309
_cons	-1.250568	19.09256	-0.07	0.948	-38.6713	36.17017

Figure 15 (Contd.): Random effects components for full model with all explanatory variables

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Identity				
var(_cons)	1.076254	.296288	.6274555	1.846065
GVKey: Independent				
var(Cen_En~s)	4793.087	1651.125	2440.018	9415.375
var(_cons)	.9610721	.0963165	.7896787	1.169665
var(Residual)	1.039075	.0343951	.9738018	1.108723

LR test vs. linear model: chi2(3) = 1822.03 Prob > chi2 = 0.0000

But there was also a drastic reduction in observation numbers due to default row-wise elimination of missing values in Stata software package. That is also probably the reason for most of the coefficients being non-significant. The only explanatory variable that was significant ($p<0.05$) was TMT role heterogeneity (positive relationship).

A summary of model-building steps from model 0 to model 4 and their progressively changing estimates is given in table 4A.

INSERT TABLE 4A ABOUT HERE

Moderation Effects of Environmental Dynamism (ED): Moderation Testing (High Dynamism) of Model 4 with All Variables, TS, Level 2 RI & RS, Level 3 RI with Mean-centered ED variable

Figure 16: Moderation under high dynamism for mean centered ED

Mixed-effects ML regression — Number of obs = 1,039

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
NAICS_Code~t	55	2	18.9	112
GVKey	312	1	3.3	6

Log likelihood = -1849.8461 — Wald chi2(10) = 110.75; Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1667872	.0215444	7.74	0.000	.124561	.2090135
Firm_Age	-.0035772	.0016951	-2.11	0.035	-.0068995	-.0002549
ln_FirmSales_CRSP_Cmpstat	.051522	.0730936	0.70	0.481	-.0917389	.1947829
Ln_Role_Hetero	.2561347	.1410644	1.82	0.069	-.0203463	.5326158
Ln_TMT_MeanTenure	-2.110677	8.236812	-0.26	0.798	-18.25453	14.03318
c.Ln_TMT_MeanTenure#c.Ln_TMT_MeanTenure	.1424079	.5425119	0.26	0.793	-.9208958	1.205712
ln_0_CEO_TotalTenure_0	-.0733837	.1875703	-0.39	0.696	-.4410147	.2942473
CEO_Pwr1_Duality_ExeCompu	-.0828882	.1227334	-0.68	0.499	-.3234412	.1576649
Log_CEO_Pwr2_ShrOwnPrcntx100	.0214505	.0414901	0.52	0.605	-.0598686	.1027696
CEO_Pwr3_FndrStat_ExeCompu	-.0730596	.3870993	-0.19	0.850	-.8317603	.6856411
_cons	9.767865	31.14128	0.31	0.754	-51.26791	70.80364

Figure 16 (Contd.)

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Identity				
var(_cons)	1.170061	.3354314	.6670948	2.052248
GVKey: Independent				
var(Cen_En~s)	351.0283	623.5576	10.79667	11412.86
var(_cons)	.965343	.1255324	.7481559	1.245579
var(Residual)	1.324014	.0705971	1.192631	1.469871

LR test vs. linear model: chi2(3) = 431.14 — Prob > chi2 = 0.0000

I tested the moderation effects by splitting the sample based on the centered mean value. All negative values represented low dynamism whereas all positive values represented high dynamism. In high dynamism scenario, once again the only significant variable (at $p<0.10$) was TMT role heterogeneity. Ceteris paribus, a very conservative interpretation may suggest that in a high dynamism situation, only TMT role heterogeneity may influence digital transformation.

Moderation Effects of Environmental Dynamism (ED): Moderation Testing (Low Dynamism) of Model 4 with All Variables, TS, Level 2 RI & RS, Level 3 RI with Mean-centered ED variable

Figure 17: Moderation under low dynamism for mean centered ED

Mixed-effects ML regression — Number of obs = 1,690

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
NAICS_Code~t	55	2	30.7	251
GVKey	315	1	5.4	7

Log likelihood = -2603.2078 — Wald chi2(10) = 192.89; Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1401379	.0139289	10.06	0.000	.1128377	.1674382
Firm_Age	-.0042179	.0016094	-2.62	0.009	-.0073724	-.0010635
ln_FirmSales_CRSP_Cmpstat	.0868405	.0621531	1.40	0.162	-.0349774	.2086584
Ln_Role_Hetero	.1436229	.1150268	1.25	0.212	-.0818255	.3690714
Ln_TMT_MeanTenure	11.63867	6.877367	1.69	0.091	-1.840726	25.11806
c.Ln_TMT_MeanTenure#c.Ln_TMT_MeanTenure	-.7606892	.4511364	-1.69	0.092	-1.6449	.123522
ln_0_CEO_TotalTenure_0	-.0935995	.1252388	-0.75	0.455	-.3390631	.1518642
CEO_Pwr1_Duality_ExeCompu	-.0496985	.0813399	-0.61	0.541	-.2091218	.1097247
Log_CEO_Pwr2_ShrOwnPrcntx100	-.0050054	.0281345	-0.18	0.859	-.0601479	.0501371
CEO_Pwr3_FndrStat_ExeCompu	-.5799999	.3134778	-1.85	0.064	-1.194405	.0344052
_cons	-41.95695	26.17792	-1.60	0.109	-93.26474	9.35084

Figure 17: Moderation under low dynamism for mean centered ED (Contd.)

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Identity				
var(_cons)	.9527474	.2868036	.5281318	1.718752
GVKey: Independent				
var(Cen_En~s)	1351.865	1232.238	226.4931	8068.853
var(_cons)	1.147546	.1175344	.9388322	1.402659
var(Residual)	.8116291	.0322652	.7507914	.8773965

LR test vs. linear model: chi2(3) = 1256.61 — Prob > chi2 = 0.0000

When I tested the low dynamism scenario (figure 17), TMT role heterogeneity no longer remained significant. But TMT tenure and TMT tenure squared, both these variables turned up significant at 0.10 level. Interestingly, the coefficients and their signs suggested that TMT tenure is positively associated with digital transformation, however beyond a point, this relationship

turns negative. So as TMT tenure increases, digital transformation may increase. But at the inflection point, any further increase in TMT tenure may lead to lower association with digital transformation. Another significant coefficient at p>0.10 was CEO founder status.

For robustness, I also used a different form of the environmental dynamism (binary, where high = 1, low = 0).

Moderation Effects of Environmental Dynamism (ED): Moderation Testing (High Dynamism) of Model 4 with All Variables, TS, Level 2 RI & RS, Level 3 RI with Binary ED variable

One advantage of using binary form of environmental dynamism was that this variable was created using a median split, so the number of observations in both, high as well as low dynamism categories were almost equal. In a high dynamism scenario, TMT role heterogeneity once again became significant at p<0.05 level, with signs remaining consistent. Founder status of CEO was also significant at p<0.05 level with a coefficient of -0.668. This being a binary variable, meant that founder CEOs (coded as 1) are negatively associated with digital transformation.

Figure 18: Moderation Effects under high dynamism (Binary ED)

Mixed-effects ML regression — Number of obs = 1,325

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
NAICS_Code~t	44	5	30.1	142
GVKey	153	2	8.7	10

Log likelihood = -2066.8771 — Wald chi2(10) = 261.67; Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1697725	.0153799	11.04	0.000	.1396285	.1999165
Firm_Age	-.0049698	.0023726	-2.09	0.036	-.00962	-.0003197
ln_FirmSales_CRSP_Cmpstat	.0174744	.0810132	0.22	0.829	-.1413086	.1762573
Ln_Role_Hetero	.307329	.1355734	2.27	0.023	.04161	.573048
Ln_TMT_MeanTenure	-5.129632	6.829248	-0.75	0.453	-18.51471	8.255449
c.Ln_TMT_MeanTenure#c.Ln_TMT_MeanTenure	.3356435	.4494366	0.75	0.455	-.545236	1.216523
ln_0_CEO_TotalTenure_0	.0639879	.1540893	0.42	0.678	-.2380215	.3659973
CEO_Pwr1_Duality_ExeCompu	-.1058452	.0965913	-1.10	0.273	-.2951607	.0834704
Log_CEO_Pwr2_ShrOwnPrcntx100	-.0026754	.0324609	-0.08	0.934	-.0662975	.0609467
CEO_Pwr3_FndrStat_ExeCompu	-.6684668	.315871	-2.12	0.034	-1.287563	-.049371
_cons	20.61261	25.90489	0.80	0.426	-30.16005	71.38526

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Identity				
var(_cons)	1.177211	.3707072	.6350507	2.182228
GVKey: Independent				
var(Cen_En~s)	2502.795	918.7079	1218.912	5138.995
var(_cons)	.9405714	.1439518	.6968154	1.269597
var(Residual)	.8896902	.0417422	.811526	.975383

LR test vs. linear model: chi2(3) = 946.49 — Prob > chi2 = 0.0000

Moderation Effects of Environmental Dynamism (ED): Moderation Testing (Low Dynamism) of Model 4 with All Variables, TS, Level 2 RI & RS, Level 3 RI with Binary ED variable

None of the variables in a low dynamism context were significant when a binary classification was used.

Figure 19: Moderation Effects under low dynamism (Binary ED)

Mixed-effects ML regression — Number of obs = 1,404

Group Variable	No. of Groups	Observations per Group Minimum	Average	Maximum
NAICS_Code~t	11	6	127.6	329
GVKey	162	2	8.7	10

Log likelihood = -2308.981 — Wald chi2(10) = 166.01; Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1434765	.0161983	8.86	0.000	.1117285	.1752245
Firm_Age	-.0030547	.0018858	-1.62	0.105	-.0067509	.0006415
ln_FirmSales_CRSP_Cmpstat	.1009201	.0756684	1.33	0.182	-.0473872	.2492275
Ln_Role_Hetero	.1273597	.150608	0.85	0.398	-.1678265	.4225459
Ln_TMT_MeanTenure	4.893426	7.347638	0.67	0.505	-9.50768	19.29453
c.Ln_TMT_MeanTenure#c.Ln_TMT_MeanTenure	-.3013365	.4825202	-0.62	0.532	-1.247059	.6443856
ln_0_CEO_TotalTenure_0	-.1469203	.1612947	-0.91	0.362	-.463052	.1692114
CEO_Pwr1_Duality_ExeCompu	-.0432951	.1019547	-0.42	0.671	-.2431227	.1565325
Log_CEO_Pwr2_ShrOwnPrcntx100	.0345033	.0360012	0.96	0.338	-.0360578	.1050643
CEO_Pwr3_FndrStat_ExeCompu	-.1404925	.3935708	-0.36	0.721	-.9118772	.6308922
_cons	-17.04462	27.92082	-0.61	0.542	-71.76843	37.67918

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Identity				
var(_cons)	.7351191	.390947	.2592247	2.084678
GVKey: Independent				
var(Cen_En~s)	18103.18	4743.268	10832.51	30253.86
var(_cons)	.9744058	.1301309	.7500026	1.265951
var(Residual)	1.132256	.0482883	1.04146	1.230967

LR test vs. linear model: chi2(3) = 840.64 — Prob > chi2 = 0.0000

A summary of different moderation effect tests for the 3-level model is given in table 5A. As several of the direct effects in the preceding models were non-significant, I decided to perform supplementary analyses.

INSERT TABLE 5A ABOUT HERE

SUPPLEMENTARY ANALYSIS: REDUCING FROM 3 LEVEL TO 2 LEVEL MODEL

The rationale for creating a 3-tier multilevel model is that the yearly observations are considered level 1, the company level information is considered level 2 and the industry level information is considered as level 3. The values of environmental dynamism are such that at the industry level, they take one single constant value across all the countries and years. However, at yearly level these values changed over time, such for a given time t, all companies within the same industry had the same environmental dynamism values. So, an alternate way of specifying the model is to fix it as a 2-level hierarchical model where each observation for each company in a given year becomes the level 1 of the model and a second level comprises of the industry-wise environmental dynamism over time.

Comparison: Model 1 with Fixed Intercept, Time Slope and Level 2 Random Intercept: Industry

Figure 20: 2-level hierarchical null model

```
Mixed-effects REML regression                   Number of obs      =      3,643
Group variable: NAICS_Code_3~t                  Number of groups   =         65

                                                Obs per group:
                                                              min =          9
                                                              avg =       56.0
                                                              max =        419

                                                Wald chi2(1)       =     418.63
Log restricted-likelihood = -6581.2133          Prob > chi2        =     0.0000
```

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1691816	.0082687	20.46	0.000	.1529753	.1853879
_cons	2.421056	.1692996	14.30	0.000	2.089235	2.752878

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Identity				
var(_cons)	1.62584	.3094697	1.119583	2.361018
var(Residual)	2.045658	.0483823	1.952995	2.142718

Estimation: Model 2 with Control Variables, Time Slope, Level 2 RI & RS (Industry)

The 2-level null model (figure 20) was followed up with a model with control variables at the first level and a random slope at the 2nd level (industry) as seen in figure 21. As evident from the figures (20 versus 20), the LR likelihood values in the subsequent model are higher (-6398 as compared to -6581), indicating better fit.

Figure 21: 2-level model with Control Variables, Time Slope, Level 2 RI & RS (Industry)

Mixed-effects REML regression — Number of obs = 3,554
Group variable: NAICS_Code_3~t — Number of groups = 65

Obs per group:
min = 7
avg = 54.7
max = 410

Wald chi2(3) = 387.47
Log restricted-likelihood = -6398.6492 — Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1656594	.0096062	17.25	0.000	.1468316	.1844871
Firm_Age	-.0026271	.0005937	-4.43	0.000	-.0037907	-.0014635
ln_FirmSales_CRSP_Cmpstat	.1533522	.022635	6.78	0.000	.1089885	.197716
_cons	1.231305	.2569768	4.79	0.000	.72764	1.734971

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Independent				
var(Cen_En~s)	488.8741	315.0706	138.2333	1728.945
var(_cons)	1.535416	.2931084	1.056168	2.232127
var(Residual)	1.996496	.0482427	1.904146	2.093325

LR test vs. linear model: chi2(2) = 1274.62 — Prob > chi2 = 0.0000

With the control variables thus identified, I proceeded to include the full model incorporating the direct effects (fixed and random at 2 levels) as hypothesized.

Full Model with all explanatory variables (Model 3 with All Variables, Time Slope, Level 2 RI & RS - Industry)

As can be seen in figure 22, the stepwise addition of model parameters improved the fit and the final direct relationships were estimated. The full model for direct relationships shows that TMT tenure (both terms), CEO duality, CEO share-owning percentage, all had significant coefficients.

Figure 22: 2-level Model with All Variables, Time Slope, Level 2 RI & RS (Industry)

Mixed-effects ML regression — Number of obs = 2,729
Group variable: **NAICS_Code_3~t** — Number of groups = 55

Obs per group:
min = 5
avg = 49.6
max = 329

Wald chi2(10) = 335.51
Log likelihood = **-4848.6862** — Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1637191	.011338	14.44	0.000	.1414971	.1859411
Firm_Age	-.0027494	.0007051	-3.90	0.000	-.0041314	-.0013674
ln_FirmSales_CRSP_Cmpstat	.1143043	.0328778	3.48	0.001	.0498651	.1787436
Ln_Role_Hetero	.0674743	.0606636	1.11	0.266	-.0514242	.1863728
Ln_TMT_MeanTenure	9.352709	4.778618	1.96	0.050	-.0132104	18.71863
c.Ln_TMT_MeanTenure#c.Ln_TMT_MeanTenure	-.6332204	.3125675	-2.03	0.043	-1.245841	-.0205994
ln_0_CEO_TotalTenure_0	-.1352441	.0857937	-1.58	0.115	-.3033966	.0329084
CEO_Pwr1_Duality_ExeCompu	-.1041932	.0626556	-1.66	0.096	-.2269958	.0186095
Log_CEO_Pwr2_ShrOwnPrcntx100	.0499736	.0224805	2.22	0.026	.0059127	.0940346
CEO_Pwr3_FndrStat_ExeCompu	-.0920115	.1832555	-0.50	0.616	-.4511858	.2671627
_cons	-32.02161	18.20528	-1.76	0.079	-67.70329	3.660081

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Independent				
var(Cen_En~s)	135.495	157.3834	13.9064	1320.175
var(_cons)	1.580549	.3262085	1.0547	2.368574
var(Residual)	1.913787	.0525687	1.813479	2.019644

LR test vs. linear model: chi2(2) = **931.96** — Prob > chi2 = **0.0000**

For the overall available sample, every unit change in TMT Tenure (natural log form) is initially associated with an increasingly positive relationship with digital transformation (natural log of frequency to word count ratio × 100000). However, this relationship becomes less strong and beyond a point (at mean tenure = 260 days), the relationship becomes marginally negative, but remains significant. CEO tenure was non-significant.

CEO Power measured using CEO duality is a binary term. The coefficient for this relationship is negative and significant at $p<0.1$ level. So, CEOs holding Chair positions are less likely to be associated with digital transformation. CEO power measured using CEO share-owning percentage is positively associated with digital transformation. Every unit change in share-owning (measured as the log of percentage × 100) is related to a 0.049 units in digital transformation (measured as log of frequency to word count ratio × 100000). So in simplified terms, every 1% increase in share-owning is associated with a 0.00000049 (4.9×10^{-7}) units increase in the digital transformation index.

A summary of model building steps for the 2-level model and their progressively changing estimates is given in table 4B.

INSERT TABLE 4B ABOUT HERE

Moderation effects of environmental dynamism (ED) : Moderation Testing (High Dynamism) of Model 3 with All Variables, TS, Level 2 RI & RS (Industry) with Mean-centered ED variable

To test my moderation effects, I split my environmental dynamism variable into high and low based on its mean centered values. All the positive terms thus became my high dynamism and all negative values became the low dynamism variable. I ran the same model for both the high and low dynamism observations (1039 for high and 1690 for low). For the high dynamism observations, none of the coefficients were significant. This might have been due to the reduction in number of observations. Alternately it could also indicate that in a high dynamism environment, none of the independent variables in my model have any effect on digital transformation. In other words, in a high dynamism environment, the firm led by its CEO and TMT members just reacts to everything happening around itself without any volition on the part of the leadership.

Figure 23: Moderation effect for high dynamism (mean-centered environmental dynamism) in 2-level model

Mixed-effects ML regression — Number of obs = 1,039
Group variable: **NAICS_Code_3~t** — Number of groups = 55

Obs per group:
min = 2
avg = 18.9
max = 112

Wald chi2(10) = 82.70
Log likelihood = **-1935.6698** — Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1822996	.0247441	7.37	0.000	.133802	.2307971
Firm_Age	-.0017166	.0011871	-1.45	0.148	-.0040433	.0006101
ln_FirmSales_CRSP_Cmpstat	.0411255	.0553293	0.74	0.457	-.067318	.149569
Ln_Role_Hetero	.1732469	.1062598	1.63	0.103	-.0350184	.3815123
Ln_TMT_MeanTenure	-.2825075	7.674911	-0.04	0.971	-15.32506	14.76004
c.Ln_TMT_MeanTenure#c.Ln_TMT_MeanTenure	-.0034664	.5041453	-0.01	0.995	-.9915731	.9846403
ln_0_CEO_TotalTenure_0	-.0386692	.1531902	-0.25	0.801	-.3389164	.261578
CEO_Pwr1_Duality_ExeCompu	-.1976786	.1090233	-1.81	0.070	-.4113604	.0160032
Log_CEO_Pwr2_ShrOwnPrcntx100	.0397239	.0382864	1.04	0.299	-.035316	.1147638
CEO_Pwr3_FndrStat_ExeCompu	.0823288	.2861982	0.29	0.774	-.4786094	.643267
_cons	4.203357	29.07174	0.14	0.885	-52.7762	61.18291

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Independent				
var(Cen_En~s)	156.9533	539.9123	.1852143	133004.5
var(_cons)	1.546863	.3782681	.9578557	2.498065
var(Residual)	2.159722	.0976606	1.976548	2.359872

LR test vs. linear model: chi2(2) = 259.50 — Prob > chi2 = 0.0000

Moderation effects of environmental dynamism (ED) : Moderation Testing (Low Dynamism) of Model 3 with All Variables, TS, Level 2 RI & RS (Industry) with Mean-centered ED variable

When I ran the model with low dynamism observations, TMT tenure showed significant coefficients for both, its linear and squared terms. So, in a low dynamism environment, every

unit of increasing rate in TMT tenure's relationship (natural log form) is positively associated with increase in digital transformation (natural log of frequency to word count ratio × 100000).

Figure 24: Moderation effect for low dynamism (mean-centered environmental dynamism) in 2-level model

Mixed-effects ML regression — Number of obs = 1,690
Group variable: NAICS_Code_3~t — Number of groups = 55

Obs per group:
min = 2
avg = 30.7
max = 251

Wald chi2(10) = 148.28
Log likelihood = -2956.795 — Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1454015	.0180504	8.06	0.000	.1100234	.1807797
Firm_Age	-.0038837	.0008669	-4.48	0.000	-.0055828	-.0021847
ln_FirmSales_CRSP_Cmpstat	.174353	.040494	4.31	0.000	.0949862	.2537198
Ln_Role_Hetero	.0117733	.0735899	0.16	0.873	-.1324603	.1560069
Ln_TMT_MeanTenure	20.03966	6.482996	3.09	0.002	7.333221	32.7461
c.Ln_TMT_MeanTenure#c.Ln_TMT_MeanTenure	-1.330598	.423106	-3.14	0.002	-2.15987	-.501325
ln_0_CEO_TotalTenure_0	-.1756803	.1044763	-1.68	0.093	-.3804501	.0290895
CEO_Pwr1_Duality_ExeCompu	-.0609269	.076946	-0.79	0.428	-.2117384	.0898846
Log_CEO_Pwr2_ShrOwnPrcntx100	.059343	.0277982	2.13	0.033	.0048595	.1138264
CEO_Pwr3_FndrStat_ExeCompu	-.2809185	.2442046	-1.15	0.250	-.7595507	.1977138
_cons	-72.79398	24.77989	-2.94	0.003	-121.3617	-24.22629

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Independent				
var(Cen_En~s)	82.04934	588.3091	.0000647	1.04e+08
var(_cons)	1.474285	.3228075	.9598499	2.264433
var(Residual)	1.773404	.0622938	1.655419	1.899799

LR test vs. linear model: chi2(2) = 549.44 — Prob > chi2 = 0.0000

But this changed to negative after the mean tenure reached 348 days (squared term). Thus the effect of moderation due to low dynamism is evident with regard to TMT mean tenure. The intensity of the relationship increased in a low dynamism environment. CEO tenure was significant at $p<0.1$ level and negative, as hypothesized. CEO duality no longer remained

significant; CEO share-owning remained significant and CEO founder status was non-significant. In a low dynamism scenario, every unit change in CEO share-owning variable was associated with 0.059 units change in the digital transformation variable. In simple terms, every 1% change in CEO share-owning was associated with an increase of 0.00000054 (5.4×10^{-7}) points on the digital transformation index.

ADDITIONAL SUPPLEMENTARY ANALYSIS FOR MODERATION (ENVIRONMENTAL DYNAMISM AS A BINARY VARIABLE)

I performed additional supplementary analyses for addressing the moderation effects in my hypotheses. For this I used a binary variable created using median split of my industry-wise environmental dynamism variable, designated as 1 for high and 0 for low dynamism.

Moderation effects of environmental dynamism (ED) : Moderation Testing (High Dynamism) of Model 3 with All Variables, TS, Level 2 RI & RS (Industry) with Binary ED variable (Value = 1)

Figure 25: Moderation effect for high dynamism (Binary ED) in 2-level model

Mixed-effects ML regression — Number of obs = 1,325
Group variable: NAICS_Code_3~t — Number of groups = 44

Obs per group:
min = 5
avg = 30.1
max = 142

Wald chi2(10) = 197.80
Log likelihood = -2284.9064 — Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1704755	.0158295	10.77	0.000	.1394502	.2015007
Firm_Age	-.0038819	.0011335	-3.42	0.001	-.0061036	-.0016602
ln_FirmSales_CRSP_Cmpstat	.0525733	.0479112	1.10	0.273	-.0413308	.1464775
Ln_Role_Hetero	.22493	.0835894	2.69	0.007	.0610978	.3887623
Ln_TMT_MeanTenure	1.115276	6.783864	0.16	0.869	-12.18085	14.4114
c.Ln_TMT_MeanTenure#c.Ln_TMT_MeanTenure	-.1042341	.4448638	-0.23	0.815	-.9761512	.767683
ln_0_CEO_TotalTenure_0	-.0086109	.1181134	-0.07	0.942	-.2401089	.222887
CEO_Pwr1_Duality_ExeCompu	-.0966635	.0882478	-1.10	0.273	-.2696259	.076299
Log_CEO_Pwr2_ShrOwnPrcntx100	.0346532	.029897	1.16	0.246	-.0239437	.0932502
CEO_Pwr3_FndrStat_ExeCompu	.1833038	.2782942	0.66	0.510	-.3621428	.7287505
_cons	-1.062009	25.8235	-0.04	0.967	-51.67513	49.55112

Figure 25 (Contd.)

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Independent				
var(Cen_En~s)	219.846	201.3882	36.50748	1323.9
var(_cons)	1.669038	.383191	1.064242	2.617531
var(Residual)	1.64765	.0657793	1.52364	1.781753

LR test vs. linear model: chi2(2) = 510.44 — Prob > chi2 = 0.0000

In the first supplementary analysis where I had used a mean split between centered values of environmental dynamism, my number of observations reduced to 1039, which I conjectured might have been the reason for none of the coefficients returning significant. This was partly substantiated in this current analysis (figure 25) where due to median split I got about 300 extra observations and also some significant coefficients. Role heterogeneity was significant and both forms of CEO tenure (linear and exponential) were significant. The directional signs on the coefficients for CEO tenure were also consistent with the previous version of testing for high dynamism (although earlier ones were not significant).

Moderation effects of environmental dynamism (ED) : Moderation Testing (Low Dynamism) of Model 3 with All Variables, TS, Level 2 RI & RS (Industry) with Binary ED variable (Value = 0)

Figure 26: Moderation effect for low dynamism (Binary ED) in 2-level model

Mixed-effects ML regression — Number of obs = 1,404
Group variable: NAICS_Code_3~t — Number of groups = 11

Obs per group:
min = 6
avg = 127.6
max = 329

Wald chi2(10) = 165.99
Log likelihood = -2543.2505 — Prob > chi2 = 0.0000

Log_NormalizedWordCount	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
Year_Id	.1597013	.0160186	9.97	0.000	.1283054	.1910972
Firm_Age	-.0023704	.0009226	-2.57	0.010	-.0041786	-.0005621
ln_FirmSales_CRSP_Cmpstat	.1623642	.046571	3.49	0.000	.0710867	.2536416
Ln_Role_Hetero	-.0882779	.0904582	-0.98	0.329	-.2655726	.0890168
Ln_TMT_MeanTenure	16.76152	6.739949	2.49	0.013	3.551457	29.97157
c.Ln_TMT_MeanTenure#c.Ln_TMT_MeanTenure	-1.107127	.4401325	-2.52	0.012	-1.969771	-.2444831
ln_0_CEO_TotalTenure_0	-.2065263	.1241781	-1.66	0.096	-.4499108	.0368582
CEO_Pwr1_Duality_ExeCompu	-.1523678	.0899054	-1.69	0.090	-.3285792	.0238436
Log_CEO_Pwr2_ShrOwnPrcntx100	.058285	.0339482	1.72	0.086	-.0082523	.1248224
CEO_Pwr3_FndrStat_ExeCompu	-.3082948	.2491356	-1.24	0.216	-.7965915	.180002
_cons	-60.04603	25.68173	-2.34	0.019	-110.3813	-9.710768

Random-effects Parameters	Estimate	Std. Err.	[95% Conf.	Interval]
NAICS_Code~t: Independent				
var(Cen_En~s)	2.35e-15	3.30e-14	2.75e-27	.0020127
var(_cons)	.9830798	.4595144	.393294	2.457311
var(Residual)	2.131998	.0807987	1.979374	2.296391

LR test vs. linear model: chi2(2) = 372.11 — Prob > chi2 = 0.0000

Interestingly, low dynamism related estimation (figure 26) gave significant coefficients for TMT mean tenure related terms (linear and squared). The direction of the coefficients were consistent with the previous supplementary analysis of low environmental dynamism. The value at which the direction of relationship between TMT tenure and digital transformation changes

was 369 days, which too was consistent with the estimate of previous analysis of low dynamism (348 days). CEO duality was marginally negatively related to the dependent variable (significant at $p<0.1$), which indicates that CEOs with dual roles are less likely to indulge in digital transformation. CEO share-owning was positively associated at $p<0.1$, i.e. every unit change in share-owning percent was related to 5.8×10^{-7} units change in digital transformation. A summary of the 2-level full model results with moderation testing is given in table 5B.

INSERT TABLE 5B ABOUT HERE

INTERPRETATION FOR HYPOTHESES SUPPORT

I tested my hypotheses against 2 alternative models of hierarchical nature – first one was the 3-level model, where observations for each company-year combination were modeled at level 1, firms were modeled at level 2 and industries were modeled at level 3. This is consistent with past literature of similar hierarchical nature. For instance, in a study that examines the relationship between TMT national diversity and firm performance (Nielsen & Nielsen, 2013), TMT characteristics were modeled at level 1, firm was modeled at level 2 and industry was modeled at level 3. This is also consistent with similar examples with time-related variables when modeled for hierarchical analysis (Hair Jr & Fávero, 2019). The other model was a 2-level model, with each company-year being at level 1 and industry being modeled at level-2. This too is consistent with literature where TMT characteristics over time have been modeled at level 1 and industry variables at level 2 (Nielsen & Nielsen, 2011).

Results from the 3-level model suggest that TMT role heterogeneity was positively associated with digital transformation, supporting Hypothesis 1(a). But the same result was not replicated in the 2-level model. Hypothesis 1(b) about moderation effect of environmental dynamism on TMT role heterogeneity was partly supported for both models. In the 3-level model, higher dynamism was associated with higher role heterogeneity (for both dynamism indexes, one with mean-centered environmental dynamism and other with binary variable), but the low dynamism scenario was non-significant. In the 2-level model, only one high dynamism-related estimation (the one using binary environmental dynamism) was positive and significant, whereas coefficients related to low dynamism were non-significant.

TMT mean tenure (hypotheses 2a) was non-significant in the 3-level model whereas it was significant at $p<0.05$ in the 2-level model. Both, the base term, and squared term were significant, providing support to hypothesis 2a in a general context. However, the moderation effect (hypothesis 2b) was not supported for the 3-level as well as 2-level models. Higher environmental dynamism was not associated with a stronger curvilinear relationship between TMT mean tenure and digital transformation. On the contrary, the coefficients under high dynamism were non-significant whereas those under low dynamism were higher than the coefficients in the general context.

Hypothesis 3 proposes a positive relationship of CEO power with digital transformation. I used three operationalizations out of the few that are consistent with extant literature - CEO duality, CEO share-owning percentage and CEO founder status. I tested these 3 operationalizations and obtained mixed results. With my 3-level model, none of the operationalizations were significant. With the 2-level model, CEO duality was negatively

associated, CEO share-holding was positively associated, and CEO founder status was non-significant against digital transformation.

Testing of Hypothesis 4a yielded coefficients in the expected direction (negative) in both the 3-level and 2-level models, however they were non-significant in both. Hypothesis 4b proposed a moderated relationship to the negative relationship such that higher environmental dynamism was expected to attenuate the negative relationship between CEO tenure and digital transformation, while low dynamism was expected to accentuate it. Testing of the relationship under high dynamism also yielded coefficients in the expected direction for both 3-level and 2-level models and within them, for both mean-centered environmental dynamism and binary environmental dynamism variables; yet all of these were non-significant. However, testing for low environmental dynamism provided significant negative coefficients across all the models, as predicted. Thus, hypothesis 4b was partially supported.

Discussion

One of the chief objectives of this study was to test and extend the Rajagopalan-Spreitzer model of strategic change to incorporate the context of digital transformation. The original model proposes that external and internal contexts both influence the way top managers make decisions and take actions towards changes in strategy. Digital transformation is a form of strategic change because a firm undergoing digital transformation experiences a "change in form, quality and/or state over time" while attempting to align with its external environment (van de Ven & Poole, 1995). To create a robust study, I first initiated arguments on why digital transformation is a form of strategic change. I argued about the terms strategy, change, organizational change and

how a combination of these terms with technological innovations manifested into strategic change. Next, I discussed the role of top management in driving several outcomes for the firms. I discussed the nature of their characteristics and the manner in which their participation influenced the firm's achievement of its objectives. I then segregated some of those characteristics of top management and CEO to identify the magnitude of influence between these two entities separately. These set of characteristics formed the internal context of the firm, as conceptualized by the Rajagopalan-Spreitzer model. The external context of the model was represented by environmental dynamism in my study. A supplementary objective of my study was to develop a measure for digital transformation. I decided to use annual reports for this purpose. This publicly available document summarizes all the key actions and strategies that a firm has taken in the past year and also discusses future plans. It thus provides a strong proxy to the firm's strategies (achieved and proposed), that in turn proxies the decision-making of the top management teams. Using an elaborate process of content analysis through text mining, and its validation through alternate channels, I created my measure for digital transformation. With all the different elements (external context, internal context, top managers' cognitions and actions, strategic change) of the Rajagopalan-Spreitzer model thus created, I proceeded to analyze my data and test my hypotheses.

As discussed earlier, I achieved mixed results. At the broad level, certain top management characteristics like team heterogeneity, team tenure, CEO tenure and CEO power were indeed associated with digital transformation. However, not all my directional hypotheses were consistent with results. TMT team heterogeneity as expected, did display a positive relationship with digital transformation. This is consistent with past literature related to other forms of strategic change. However, one difference between the conceptualization of the

construct of TMT heterogeneity in this study was the heterogeneity of roles that the TMT member has worked in within the same organization. So, TMT members within an organization who have had wider or broader portfolios during their tenure in the organization will positively influence digital transformation. However, one thing to note here is that I tested 2 different forms of hierarchy – 2-level and 3-level, out of which this variable was significant only in the 3-level hierarchy. One explanation could be the idiosyncrasies associated with data and the manner in which data are structured.

TMT mean tenure showed a curvilinear relationship with digital transformation in the 2-level hierarchical model, but was not significant in the 3-level model. In a generic context, there is an initial positive relationship - shorter tenure is associated with lower quantum and as mean tenure grows, so does the quantum of digital transformation. So longer the TMT spends time together, higher is its influence on strategic change. But this relationship does not remain consistent over time. Beyond a mean tenure value of approximately 261 days, this positive relationship peaks, after which there is a negative trend. While my current data show this to be the peak number, it is likely that there may be unique differences among companies, industries, and actual TMT members. The current result was significant only for the 2-level model; hence it could not be estimated for company differences. It must have been due to a large number of observations being dropped by the estimating software. Therefore, examination of such differences may form a part of set of future research questions. Moderation testing of these relationships as per hypotheses also did not yield wholistic results. However, at a macro level, the partial results indicate that the curvilinear relationship is still retained under specific conditions (in this case low environmental dynamism). One way to interpret this may also be that at high environmental dynamism, the relationship between TMT mean tenure becomes

redundant, i.e. it does not matter whether the tenure is long or short. The pressure due to high environmental dynamism forces the top management teams to react to those frequent dynamic changes without any influence of the internal characteristics. This interpretation could be applied within the context of the dichotomy between deliberate versus emergent strategies (Mintzberg & Waters, 1985). Past literature has provided empirical support to both these perspectives and my result may provide the boundary condition under which each of these will be applicable. When the environmental dynamism is high, firms may manifest emergent strategies in which TMT characteristics like team tenure may not matter at all, hence the non-significant coefficients. Continuing this argument, it may also explain significant and higher coefficients for team tenure and the squared team tenure variable under low dynamism. Under this situation, TMT tenure may continue to influence digital transformation because without the environmental pressure, the top executives may find it more productive to work on implementing newer digital transformation strategies.

I used three different measures of CEO power for hypothesis 3. CEO Duality is a form of structural power whereas share-owning percentage and founder status are forms of ownership power. All three forms of CEO power were expected to positively influence digital transformation. However, my results were not consistent with the expectations. CEO duality, a binary variable with 1 representing dual responsibility of CEO, showed a negative significant coefficient ($p<0.1$) in the 2-level model and was non-significant in the 3-level model. Some previous studies have found a negative relationship (Berg & Smith, 1978) between duality and firm performance whereas some have found no significant results (Boyd, 1995; Chaganti, Rajeswararao S., Mahajan, & Sharma, 1985; Rechner & Dalton, 1991). All these studies have reported certain contingencies under which these results are observed. Berg and Smith found that

the negative relationship existed in certain industries, whereas Rechner and Dalton found that the negative relationship was observed only in some operationalizations of firm performance (RoE, RoI and profit margins). More detailed analysis of my data may reveal the exact reason for these negative results.

CEO ownership power may determine the efficacy of implementing decisions of his/ her choice (Daily & Johnson, 1997). So implementation of digital transformation may be a function of ownership power. My analysis using a 2-level model obtained a positive significant coefficient for CEO share-owning, i.e. higher percentage of share ownership by CEO may determine higher digital transformation scores. However, the coefficients were not significant for founder status, another measure of ownership power. This could be due to the very small percentage of founder - CEOs in my dataset. A supplementary analysis with a study of mostly founder-driven firms may provide different results.

CEO tenure was expected to be negatively associated with digital transformation. My results did not provide support for the hypothesized direct relationship. Past research on CEO performance has indeed found a negative relationship between CEO tenure and attitude towards change (Musteen, Barker III, & Baeten, 2006) and firm performance. However, under certain contingencies, for instance industry characteristics, the relationship may manifest differently (Henderson, Miller, & Hambrick, 2006). These contingencies were hypothesized as moderations. None were supported in the 3-level model. However, in the 2-level model, the low dynamism scenario had negative significant coefficients, thus offering partial support. Low dynamism scenarios were associated with negative relationship, whereas coefficients in high dynamism scenario were non-significant. Another way to interpret the moderation hypothesis is to associate attenuated relationship with non-significance. If a relationship is negative in one scenario but not

significant another, it may also mean that no relationship exists. In other words, the effect of dynamism on the relationship between CEO tenure and digital transformation exists only in low dynamism scenarios. So in the event of attenuation of the negative relationship, it gets so diminished that it no longer affects the phenomenon.

Overview of the Results and Discussion for Theory Enhancement

The phenomenon of digital transformation studied as strategic change in this dissertation is an attempt to create an extension of the Rajagopalan-Spreitzer model of strategic change. With the results obtained in this study, some elements of extension to the model are being proposed. It is a very well developed and wholistic model incorporating the relationship between external environment, internal environment, and managerial cognitions and actions that determine changes to strategy and therefore performance. Yet, this model is static in nature. It incorporates changes to these elements and how those changes eventually affect strategic decision-making, but it does not incorporate the element of time in the model. It provides a framework to study strategic change at a cross-section, but not over a period of time. One basic characteristic of change, as expressed in the seminal definition of van de Ven and Poole, is the "change in structure, form or quality of firm over ***time"*** (emphasis added). So, the element of time is required to be added. Secondly, managerial cognitions are a function of the unique characteristics that each individual executive possesses and the consolidated influence of the interactions of these unique characteristics. In the present model, environmental and organizational conditions are directly shown to impact these managerial cognitions and actions. It is therefore an incomplete manifestation of the actual way in which these relationships may

work. Inclusion of a term representing managerial characteristics may address this inconsistency. Another set of elements to be incorporated is feedback loops. These are essential to accommodate the technology-enabled strategic changes like digital transformation, which enable real-time feedback to the firm at every moment of time. While previously, any change and performance effects therein were available only after some time lags, the digital readiness of firms enable them to be equipped to monitor every strategic move on real-time basis. This study can provide a basic framework to add these elements into the existing model.

Methodological Considerations

There are increasing calls from contemporary scholars to integrate micro and macro domains of management. For instance, in a special issue on this subject in the Journal of Management, the editors (Aguinis, Boyd, Pierce, & Short, 2011) lay out a summary of the general trends towards initiating the movement to bridge this gap. But one most basic underlying sentiment expressed by every author in this special issue is the development of newer paradigms on multilevel research (Dalton & Dalton, 2011; De Vries, Hollenbeck, Davison, Walter, & Van Der Vegt, Gerben S, 2016; Huselid & Becker, 2011; Molina-Azorín, 2014; Priem, Walters, & Li, 2011). While every macro level phenomenon can be modeled as a multilevel problem, the complexities involved in consolidating these different level variables preclude many scholars from pursuing this route. One of the management domains that can be more readily built into multilevel models, is TMT research. Some scholars have already initiated this movement (Hair Jr & Fávero, 2019; Nielsen & Nielsen, 2011; Nielsen & Nielsen, 2013; Short, Palmer, & Ketchen, 2003; Short, Ketchen Jr, Palmer, & Hult, 2007; Steenbergen & Jones, 2002). I have also

highlighted earlier the conceptualization of strategic change as a multilevel model problem by integrating it with the strategic human capital resources approach (Ployhart & Moliterno, 2011; Ployhart, Nyberg, Reilly, & Maltarich, 2014), the upper echelons approach (Bromiley & Rau, 2016; Carpenter et al., 2004; Hambrick, 2007), and the dynamic managerial capabilities approach (Helfat et al., 2009; Helfat & Martin, 2015; Teece, Pisano, & Shuen, 1997; Teece, 2007). With the inclusion of individual level variables that change with time within the organization, which is nested in an industry, this study is an appropriate avenue. Through this study, I have made an attempt to demonstrate the advantage of a multilevel approach - less biased estimates due to controlling out the random effects of company and industry on my variables of interest. This takes me to the next part of my dissertation.

Limitations and future research

This study started as an ambitious attempt to consolidate the phenomenon of strategic change and integrate it with the new technology perspectives, notably digital transformation. While strategic change by itself is a well-developed sub-domain transcending the larger strategy, management and organizations domains, digital transformation as of yet is a very recent phenomenon. This itself determines the first big limitation of this research – liability of newness. Academic understanding about digital transformation is still evolving. During the process of trying to consolidate my understanding about this phenomenon, my conversations with several subject matter experts indicated that it is yet a very under-defined and under-developed domain. The large umbrella term called “digital transformation” was understood in multiple different ways by multiple experts. Hence it was essential to carry out a detailed pre- and post-analysis of

my process of variable creation related to digital transformation. To consolidate this wide variance in understanding, I used three different avenues to define the domain. One was the publicly available popular literature searchable through google. I identified 14 websites (belonging to key organizations that are popularly considered to be leaders in digital transformation) which contained some kind of glossary of terms related to digital transformation. This effort provided me with a list of about 250 key terms associated with the phenomenon. The second avenue was to carry out text analysis on all the extant academic literature related to digital transformation and extract the top 220 frequent occurring terms from academic literature. The next step was to carry out a text analysis of 5 top-rated books on the subject and extract the 150 most frequently occurring key-terms used in those books. These three lists were consolidated using a process of elimination in which key-terms occurring in atleast 2 of these sources were first isolated, giving a list of about 460 key terms. This list was then sent to few subject matter experts for their validation. Final set of responses from 8 of them were used to create the custom dictionary that represented the domain of digital transformation. After eliminating all the terms that received lower than average rating of 1.75 were eliminated and the shortlisted terms were sent to 2 academic scholars for their vetting, finally giving a list of 348 terms associated with digital transformation.

One criticism of this approach may emerge from an argument about validity of the measure, i.e., whether a list of words actually represent strategic actions that a firm is taking or has taken. I have attempted to address this limitation in an earlier section that discusses salient features of content analysis. A combination of arguments related to importance of annual reports in reflecting managerial cognitions (Bowman, 1978; Bowman, 1984), and a causal mapping

between managerial cognitions through annual reports and strategic actions (Barr et al., 1992; Nadkarni & Barr, 2008) may suitably address this criticism.

Another limitation of this study is that it tested only a limited number of variables related to top management teams. One reason was to limit the scope of the project. For instance, several other variables like TMT functional heterogeneity, TMT educational background, CEO functional experience, CEO personality, etc. were eventually dropped from the initially proposed list. Another reason was the lack of availability of data associated with some of the variables. For instance, my data for the different variables were sourced from different databases like Compustat, ExecuComp, BoardEx, and Corporate Affiliations. My data for the dependent variable were sourced from annual reports extracted from the annualreports.com website and respective companies for few of them. Striking the most optimal balance that included maximum common observations from all these sources drastically reduced my number of observations from 3600 to about 2700. This drastic reduction may have contributed to the several non-significant results that I obtained. Future research will examine the influence of some of the TMT and CEO characteristics that were not included in this study. Another extension of this study is to propose and test interactions between the variables of interest. In the present study, this has not been done. A third extension of this work could also be the inclusion of board characteristics in addition to TMT and CEO characteristics.

One more unintended consequence of including a very limited number of variables, is omitted variable bias, which in turn may give rise to endogeneity issues. Testing and correcting for endogeneity involves identification of instrumental variables and their theoretical justifications. This approach is proposed to be undertaken in the near future.

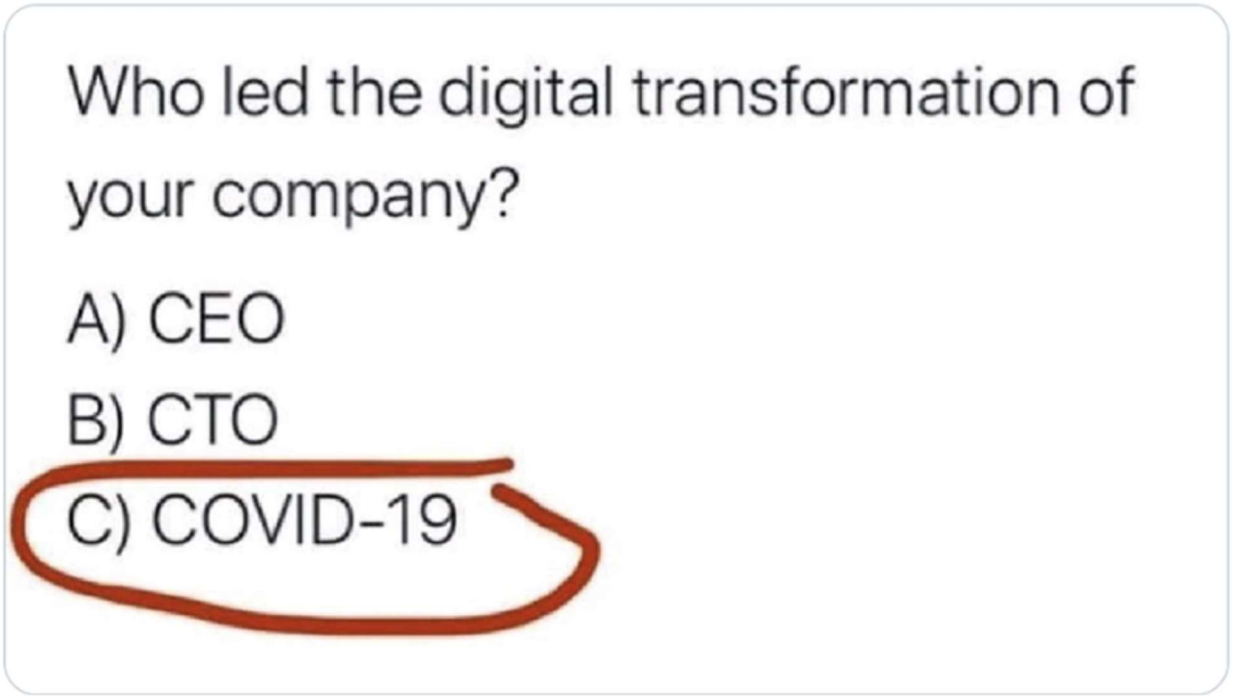
Many of the hypothesized relationships were not found significant for which several probable reasons have already been discussed earlier. As a result, it is difficult to create a wholistic perspective about the relationships that were examined. For instance, out of the 2 variables related to TMT, one variable provided support in one model whereas another variable provided support in another model. This is like comparing apples to oranges. Future endeavors will need to incorporate these challenges.

CONCLUSION

This study attempts to extend the multi-lens model of strategic change to incorporate this new way in which organizations are experiencing change. Until just four months back, digital transformation was a buzzword fashionably uttered in several boardrooms. But one natural event has altered the course of manner in which the whole world views digital transformation today. The Covid19 pandemic has exponentially increased the pace at which organizations are gearing up for and implementing multiple initiatives. A meme featured in one of the blogs goes like this:

(https://blog.smarp.com/how-covid-19-is-accelerating-digital-transformation-in-the-workplace)

Figure 27: Meme featured in several digital transformation forums (Source: blog.smarp.com)



Although this is presented as a joke, a reality check on several companies which expedited their digital transformation efforts during last few months may unravel this exact story. Discussed in the context of my present dissertation, changes in external environment have forced firms to undergo digital transformation. None of the upper echelons characteristics would hold significant support in the related initiatives taken by the firms. One prospective research question would be: how have the speed and quality of digital transformation initiatives changed before, during, and after the pandemic? A supplementary question could be: how has the relationship between upper echelons characteristics and digital transformation been moderated by the pandemic?

In conclusion, the fourth industrial revolution has accelerated in the past few years and more so in the last few months. Firms that had been preparing for digital initiatives may be able to recover faster from the debilitating economic slowdown, offering them a transient competitive advantage over those firms which were not digital ready till few months back.

REFERENCES

Abdul Hameed, M., & Counsell, S. (2012). Assessing the influence of environmental and CEO characteristics for adoption of information technology in organizations. *Journal of Technology Management & Innovation, 7*(1), 64-84.

Adner, R., & Helfat, C. E. (2003). Corporate effects and dynamic managerial capabilities. *Strategic Management Journal, 24*(10), 1011-1025.

Aguinis, H., Boyd, B. K., Pierce, C. A., & Short, J. C. (2011). No title. *Walking New Avenues in Management Research Methods and Theories: Bridging Micro and Macro Domains,*

Amason, A. C., & Sapienza, H. J. (1997). The effects of top management team size and interaction norms on cognitive and affective conflict. *Journal of Management, 23*(4), 495-516.

Andersson, P., Movin, S., Mähring, M., Teigland, R., & Wennberg, K. (2018). *Managing digital transformation* SSE Institute for Research, Stockholm School of Economics.

Andrews, K. R. (1971). Concept of corporate strategy.

Arceneaux, K., & Nickerson, D. W. (2009). Modeling certainty with clustered data: A comparison of methods. *Political Analysis,* , 177-190.

Armenakis, A. A., & Bedeian, A. G. (1999). Organizational change: A review of theory and research in the 1990s. *Journal of Management, 25*(3), 293-315.

Bantel, K. A., & Jackson, S. E. (1989). Top management and innovations in banking: Does the composition of the top team make a difference? *Strategic Management Journal, 10*(S1), 107-124.

Barker, V. L., & Mueller, G. C. (2002). CEO characteristics and firm R&D spending. *Management Science, 48*(6), 782-801. doi:10.1287/mnsc.48.6.782.187

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99-120.

Barr, P. S., Stimpert, J. L., & Huff, A. S. (1992). Cognitive change, strategic action, and organizational renewal. *Strategic Management Journal, 13*(8), 15. Retrieved from https://search.proquest.com/docview/1305781958 https://search.proquest.com/docview/1305781958

Barry, E. J., Kemerer, C. F., & Slaughter, S. A. (2006). Environmental volatility, development decisions, and software volatility: A longitudinal analysis. *Management Science, 52*(3), 448-464.

Berg, S. V., & Smith, S. K. (1978). CEO and board chairman: A quantitative study of dual vs. unitary board leadership. *Directors and Boards, 3*(1), 34-39.

Boeker, W. (1997). Strategic change: The influence of managerial characteristics and organizational growth. *The Academy of Management Journal, 40*(1), 152-170. doi:10.2307/257024

Böhm, M., Leimeister, S., Riedl, C., & Krcmar, H. (2010). Cloud computing and computing evolution. *Technische Universität München (TUM), Germany,*

Bowman, E. H. (1976). Strategy and the weather. *Sloan Management Review, 17*(2), 49-62. Retrieved from https://search.proquest.com/docview/1302986072 https://search.proquest.com/docview/1302986072

Bowman, E. H. (1978). Strategy, annual reports, and alchemy. *California Management Review, 20*(3), 64-71. doi:10.2307/41165283

Bowman, E. H. (1984). Content analysis of annual reports for corporate strategy and risk. *Interfaces, 14*(1), 61-71.

Boyd, B. K. (1995). CEO duality and firm performance: A contingency model. *Strategic Management Journal, 16*(4), 301-312.

Bromiley, P., & Rau, D. (2016). Social, behavioral, and cognitive influences on upper echelons during strategy process: A literature review. *Journal of Management, 42*(1), 174-202.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies* WW Norton & Company.

Busenbark, J. R., Krause, R., Boivie, S., & Graffin, S. D. (2016). *Toward a configurational perspective on the CEO*. Los Angeles, CA: SAGE Publications. doi:10.1177/0149206315618448

Carpenter, M. A., Geletkanycz, M. A., & Sanders, W. G. (2004). Upper echelons research revisited: Antecedents, elements, and consequences of top management team composition. *Journal of Management, 30*(6), 749-778.

Certo, S. T., Lester, R. H., Dalton, C. M., & Dalton, D. R. (2006). Top management teams, strategy and financial performance: A meta-analytic examination. *Journal of Management Studies, 43*(4), 813-839.

Chaffee, E. E. (1985). Three models of strategy. *The Academy of Management Review, 10*(1), 89-98. doi:10.2307/258215

Chaganti, R. S., Mahajan, V., & Sharma, S. (1985). Corporate board size, composition and corporate failures in retailing industry [1]. *Journal of Management Studies, 22*(4), 400-417.

Chaganti, R., & Sambharya, R. (1987). Strategic orientation and characteristics of upper management. *Strategic Management Journal, 8*(4), 393-401.

Chandler, A. D. (1962). Strategy and structure: Chapters in the history of american industrial enterprises. *Cambridge.Hlass.: MIT Press, 14*, 16.

Child, J. (1972). Organizational structure, environment and performance: The role of strategic choice. *Sociology, 6*(1), 1-22.

Cho, T. S., & Hambrick, D. C. (2006). Attention as the mediator between top management team characteristics and strategic change: The case of airline deregulation. *Organization Science, 17*(4), 453-469.

Clapham, S. E., & Schwenk, C. R. (1991). Self-serving attributions, managerial cognition, and company performance. *Strategic Management Journal, 12*(3), 219-229. doi:10.1002/smj.4250120305

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly, 35*(1), 128-152.

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. doi:10.1111/j.1467-6486.2009.00880.x

Cyert, R. M., & March, J. G. (1963). A behavioral theory of the firm. *Englewood Cliffs, NJ, 2*(4), 169-187.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*(5), 554-571.

Daily, C. M., & Johnson, J. L. (1997). Sources of CEO power and firm financial performance: A longitudinal assessment. *Journal of Management, 23*(2), 97-117.

Dalton, D. R., & Dalton, C. M. (2011). No title. *Integration of Micro and Macro Studies in Governance Research: CEO Duality, Board Composition, and Financial Performance,*

Datta, D. K., Rajagopalan, N., & Zhang, Y. (2003). New CEO openness to change and strategic persistence: The moderating role of industry characteristics. *British Journal of Management, 14*(2), 101-114.

Davenport, T. H., & Harris, J. G. (2007). Competing on analytics: The new science of winning. *Harvard Business Review Press,*

De Vries, T. A., Hollenbeck, J. R., Davison, R. B., Walter, F., & Van Der Vegt, Gerben S. (2016). Managing coordination in multiteam systems: Integrating micro and macro perspectives. *Academy of Management Journal, 59*(5), 1823-1844.

Dess, G. G., & Beard, D. W. (1984). Dimensions of organizational task environments. *Administrative Science Quarterly,* , 52-73.

Dugal, M., & Gopalakrishnan, S. (2000). Environmental volatility: A reassessment of the construct. *The International Journal of Organizational Analysis, 8*(4), 401-424.

Eggers, J. P., & Kaplan, S. (2013). Cognition and capabilities: A multi-level perspective. *Academy of Management Annals, 7*(1), 295-340.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: What are they? *Strategic Management Journal, 21*(10), 1105-1121.

Ferrier, W. J. (2001). Navigating the competitive landscape: The drivers and consequences of competitive aggressiveness. *Academy of Management Journal, 44*(4), 858-877.

Finkelstein, S. (1992). Power in top management teams: Dimensions, measurement, and validation. *Academy of Management Journal, 35*(3), 505-538.

Finkelstein, S., & Hambrick, D. C. (1990). Top-management-team tenure and organizational outcomes: The moderating role of managerial discretion. *Administrative Science Quarterly,* , 484-503.

Garbuio, M., King, A. W., & Lovallo, D. (2011). Looking inside: Psychological influences on structuring a firmâ€™s portfolio of resources. *Journal of Management, 37*(5), 1444-1463.

Gioia, D. A., & Chittipeddi, K. (1991). Sensemaking and sensegiving in strategic change initiation. *Strategic Management Journal, 12*(6), 433-448. Retrieved from https://www.jstor.org/stable/2486479 https://www.jstor.org/stable/2486479

Goll, I., Brown Johnson, N., & Rasheed, A. A. (2007). Knowledge capability, strategic change, and firm performance: The moderating role of the environment. *Management Decision, 45*(2), 161-179.

Goodstein, J., Gautam, K., & Boeker, W. (1994). The effects of board size and diversity on strategic change. *Strategic Management Journal, 15*(3), 241-250.

Greenway, A., Terrett, B., Bracken, M., & Loosemore, T. (2018). *Digital transformation at scale: Why the strategy is delivery: Why the strategy is delivery* London Publishing Partnership.

Gupta, A. K. (1984). Contingency linkages between strategy and general manager characteristics: A conceptual examination. *Academy of Management Review, 9*(3), 399-412.

Hair Jr, J. F., & Fávero, L. P. (2019). Multilevel modeling for longitudinal data: Concepts and applications. *RAUSP Management Journal,*

Hambrick, D. C. (2007). Upper echelons theory: An update. *The Academy of Management Review, 32*(2), 334-343. doi:10.2307/20159303

Hambrick, D. C., Cho, T. S., & Chen, M. (1996). The influence of top management team heterogeneity on firms' competitive moves. *Administrative Science Quarterly,* , 659-684.

Hambrick, D. C., & Finkelstein, S. (1987). Managerial discretion: A bridge between polar views of organizational outcomes. *Research in Organizational Behavior,*

Hambrick, D. C., Finkelstein, S., & Mooney, A. C. (2005). Executive job demands: New insights for explaining strategic decisions and leader behaviors. *The Academy of Management Review, 30*(3), 472-491. doi:10.2307/20159139

Hambrick, D. C., & Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *The Academy of Management Review, 9*(2), 193-206. doi:10.2307/258434

Hambrick, D. C., Geletkanycz, M. A., & Fredrickson, J. W. (1993). Top executive commitment to the status quo: Some tests of its determinants. *Strategic Management Journal, 14*(6), 401-418.

Heck, R. H., Thomas, S. L., & Tabata, L. N. (2009). An introduction to multilevel modeling techniques (quantitative methodology series).

Helfat, C. E., Finkelstein, S., Mitchell, W., Peteraf, M., Singh, H., Teece, D., . . . Winter, S. G. (2009). *Dynamic capabilities: Understanding strategic change in organizations* John Wiley & Sons.

Helfat, C. E., & Martin, J. A. (2015). *Dynamic managerial capabilities*. Los Angeles, CA: SAGE Publications. doi:10.1177/0149206314561301

Helfat, C. E., & Peteraf, M. A. (2015). Managerial cognitive capabilities and the microfoundations of dynamic capabilities. *Strategic Management Journal, 36*(6), 831-850. doi:10.1002/smj.2247

Henderson, A. D., Miller, D., & Hambrick, D. C. (2006). How quickly do CEOs become obsolete? industry dynamism, CEO tenure, and company performance. *Strategic Management Journal, 27*(5), 447-460.

Herrmann, P., & Datta, D. K. (2006). CEO experiences: Effects on the choice of FDI entry mode. *Journal of Management Studies, 43*(4), 755-778.

Hinings, B., Gegenhuber, T., & Greenwood, R. (2018). Digital innovation and transformation: An institutional perspective. *Information and Organization, 28*(1), 52-61. doi:10.1016/j.infoandorg.2018.02.004

Hodgkinson, G. P., & Healey, M. P. (2011). Psychological foundations of dynamic capabilities: Reflexion and reflection in strategic management. *Strategic Management Journal, 32*(13), 1500-1516.

Hofer, C. W. (1973). (1973). Some preliminary research on patterns of strategic behavior. Paper presented at the *, 1973*(1) 46-54.

Hofer, C. W., & Schendel, D. (1978). Strategy formulation: Analytical concepts, west, st. *Paul, MN,*

Hough, J. R., & Ogilvie, D. T. (2005). An empirical test of cognitive style and strategic decision outcomes. *Journal of Management Studies, 42*(2), 417-448.

Huff, A. (1990). *Mapping strategic thought* John Wiley & Sons.

Huff, A., & Reger, R. (1987). A review of strategic process research. *Journal of Management, 13*(2), 211-236.

Huselid, M. A., & Becker, B. E. (2011). No title. *Bridging Micro and Macro Domains: Workforce Differentiation and Strategic Human Resource Management,*

Jiao, H., Alon, I., Koo, C. K., & Cui, Y. (2013). When should organizational change be implemented? the moderating effect of environmental dynamism between dynamic capabilities and new venture performance. *Journal of Engineering and Technology Management, 30*(2), 188-205. doi:10.1016/j.jengtecman.2013.01.005

Judson, A. S. (1991). *Changing behavior in organizations: Minimizing resistance to change* B. Blackwell.

Keck, S. L. (1997). Top management team structure: Differential effects by environmental context. *Organization Science, 8*(2), 143-156.

Kelly, D., & Amburgey, T. L. (1991). Organizational inertia and momentum: A dynamic model of strategic change. *Academy of Management Journal, 34*(3), 591-612.

Kerravala, Z., & Miller, L. C. (2017). *Digital transformation for dummies*. 111 River St: John Wiley & Sons, Inc.

King, W. R., Grover, V., & Hufnagel, E. H. (1989). Using information and information technology for sustainable competitive advantage: Some empirical evidence. *Information & Management, 17*(2), 87-93.

Kitchell, S. (1997). CEO characteristics and technological innovativeness: A canadian perspective. *Canadian Journal of Administrative Sciences/Revue Canadienne Des Sciences De L'Administration, 14*(2), 111-121.

Knight, D., Pearce, C. L., Smith, K. G., Olian, J. D., Sims, H. P., Smith, K. A., & Flood, P. (1999). Top management team diversity, group process, and strategic consensus. *Strategic Management Journal, 20*(5), 445-465.

Kotter, J. P. (1995). Leading change: Why transformation efforts fail (vol. 73). *Cambridge, MA: Harvard Business School Publication,*

Kozlowski, S. W. J., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest, 7*(3), 77-124.

Lefebvre, E., & Lefebvre, L. A. (1992). Firm innovativeness and CEO characteristics in small manufacturing firms. *Journal of Engineering and Technology Management, 9*(3-4), 243-277.

Lewin, K. (1947). Frontiers in group dynamics: Concept, method and reality in social science; social equilibria and social change. *Human Relations, 1*(1), 5-41.

Lewis, B. W., Walls, J. L., & Dowell, G. W. S. (2014). Difference in degrees: CEO characteristics and firm environmental disclosure. *Strategic Management Journal, 35*(5), 712-722.

Lin, C., Lin, P., Song, F. M., & Li, C. (2011). Managerial incentives, CEO characteristics and corporate innovation in chinaâ€™s private sector. *Journal of Comparative Economics, 39*(2), 176-190.

Lipman, B. L. (1995). Information processing and bounded rationality: A survey. *The Canadian Journal of Economics / Revue Canadienne D'Economique, 28*(1), 42-67. doi:10.2307/136022

Manyika, J., Chui, M., Bughin, J., Dobbs, R., Bisson, P., & Marrs, A. (2013). *Disruptive technologies: Advances that will transform life, business, and the global economy.* (). San Francisco, USA:

March, J. G., & Simon, H. A. (1958). Organizations john wiley & sons. *New York,*

Mata, F. J., Fuerst, W. L., & Barney, J. B. (1995). Information technology and sustained competitive advantage: A resource-based analysis. *MIS Quarterly,* , 487-505.

Messick, S. (1976). Personality consistencies in cognition and creativity. *Individuality in Learning, 4*, 22.

Miller, D. (1991). Stale in the saddle: CEO tenure and the match between organization and environment. *Management Science, 37*(1), 34-52.

Miller, F., Vandome, A. F., & McBrewster, J. (2010). *Amazon web services* Alpha Press.

Mintzberg, H., & Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic Management Journal, 6*(3), 257-272.

Mischel, W. (1977). The interaction of person and situation. *Personality at the Crossroads: Current Issues in Interactional Psychology, 333*, 352.

Molina-Azorín, J. F. (2014). Microfoundations of strategic management: Toward micro–macro research in the resource-based theory. *BRQ Business Research Quarterly, 17*(2), 102-114.

Musteen, M., Barker III, V. L., & Baeten, V. L. (2006). CEO attributes associated with attitude toward change: The direct and moderating effects of CEO tenure. *Journal of Business Research, 59*(5), 604-612.

Nadkarni, S., & Barr, P. S. (2008). Environmental context, managerial cognition, and strategic action: An integrated view. *Strategic Management Journal, 29*(13), 1395-1427.

Nielsen, B. B., & Nielsen, S. (2011). The role of top management team international orientation in international strategic decision-making: The choice of foreign entry mode. *Journal of World Business, 46*(2), 185-193.

Nielsen, B. B., & Nielsen, S. (2013). Top management team nationality diversity and firm performance: A multilevel study. *Strategic Management Journal, 34*(3), 373-382.

Orlikowski, W. J. (1993). CASE tools as organizational change: Investigating incremental and radical changes in systems development. *MIS Quarterly,* , 309-340.

Palgrave, R. H. I. (1987). *The new palgrave: A dictionary of economics* Macmillan.

Palmer, N., Moore, C., Finn, K. M., Winkler, K., Kowalkowski, F., & Swenson, K. D. (2019). *Digital transformation with BPM* Future Strategies Inc.

Penrose, E. T. (1959). The theory of the growth of the firm (1995).

Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resourceâ€ based view. *Strategic Management Journal, 14*(3), 179-191.

Piccarozzi, M., Aquilani, B., & Gatti, C. (2018). Industry 4.0 in management studies: A systematic literature review. *Sustainability, 10*(10), 3821. doi:10.3390/su10103821

Piccoli, G., & Ives, B. (2005). IT-dependent strategic initiatives and sustained competitive advantage: A review and synthesis of the literature. *MIS Quarterly, 29*(4), 747-776.

Ployhart, R. E., & Moliterno, T. P. (2011). Emergence of the human capital resource: A multilevel model. *Academy of Management Review, 36*(1), 127-150.

Ployhart, R. E., Nyberg, A. J., Reilly, G., & Maltarich, M. A. (2014). Human capital is dead; long live human capital resources! *Journal of Management, 40*(2), 371-398.

Porter, M. E. (1980). Competitive strategy: Techniques for analyzing industries and companies. *New York,*

Priem, R. L., Walters, B. A., & Li, S. (2011). Decisions, decisions! how judgment policy studies can integrate macro and micro domains in management research. *Journal of Management, 37*(2), 553-580.

Quigley, T. J., & Hambrick, D. C. (2015). Has the â€œCEO effectâ€ increased in recent decades? A new explanation for the great rise in america's attention to corporate leaders. *Strategic Management Journal, 36*(6), 821-830. doi:10.1002/smj.2258

Rajagopalan, N., & Spreitzer, G. M. (1997). Toward a theory of strategic change: A multi-lens perspective and integrative framework. *Academy of Management Review, 22*(1), 48-79.

Rechner, P. L., & Dalton, D. R. (1991). CEO duality and organizational performance: A longitudinal analysis. *Strategic Management Journal, 12*(2), 155-160.

Rogers, D. L. (2016). *The digital transformation playbook: Rethink your business for the digital age* Columbia University Press.

Rumelt, R. P. (1984). Towards a strategic theory of the firm. teoksessa lamb, RB (toim.): Competitive strategic management, 556-570.

Salancik, G. R., & Meindl, J. R. (1984). Corporate attributions as strategic illusions of management control. *Administrative Science Quarterly, 29*(2), 238-254. doi:10.2307/2393176

Schendel, D., & Hofer, C. W. (1979). *Strategic management: A new view of business policy and planning* Little, Brown.

Schultz, T. W. (1971). *Investment in human capital*. New York: Free Press u.a.].

Short, J. C., Ketchen Jr, D. J., Palmer, T. B., & Hult, G. T. M. (2007). Firm, strategic group, and industry influences on performance. *Strategic Management Journal, 28*(2), 147-167.

Short, J. C., Palmer, T. B., & Ketchen, D. J. (2003). Multilevel influences on firm performance: Insights from the resource-based view and strategic groups research. *Research in Multi-Level Issues, 2*(1), 155-188.

Staw, B. M., McKechnie, P. I., & Puffer, S. M. (1983). The justification of organizational performance. *Administrative Science Quarterly, 28*(4), 582-600. doi:10.2307/2393010

Steenbergen, M. R., & Jones, B. S. (2002). Modeling multilevel data structures. *American Journal of Political Science,* , 218-237.

Teece, D. J. (1982). Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior & Organization, 3*(1), 39-63.

Teece, D. J. (2007). Explicating dynamic capabilities. *Strategic Management Journal, 28*(13), 1319-1350. Retrieved from http://www.econis.eu/PPNSET?PPN=550362401 http://www.econis.eu/PPNSET?PPN=550362401

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18*(7), 509-533. doi:AID-SMJ882>3.0.CO;2-Z

Thomas, A. S., Litschert, R. J., & Ramaswamy, K. (1991). The performance impact of strategy-manager coalignment: An empirical examination. *Strategic Management Journal, 12*(7), 509-522.

Thong, J. Y. L., & Yap, C. (1995). CEO characteristics, organizational characteristics and information technology adoption in small businesses. *Omega, 23*(4), 429-442.

Thorelli, H. B. (1977). *Strategy structure: The strategic planning imperative* Indiana University Press.

van de Ven, A., & Poole, M. (1995). Explaining development and change in organizations. *The Academy of Management Review, 20*(3), 510-540. doi:10.2307/258786

Walsh, J. P. (1995). Managerial and organizational cognition: Notes from a trip down memory lane. *Organization Science, 6*(3), 280-321.

Wei, L., & Wu, L. (2013). What a diverse top management team means: Testing an integrated model. *Journal of Management Studies, 50*(3), 389-412.

Wernerfelt, B. (1984). A resourceâ€ based view of the firm. *Strategic Management Journal, 5*(2), 171-180.

West Jr, C. T., & Schwenk, C. R. (1996). Top management team strategic consensus, demographic homogeneity and firm performance: A report of resounding nonfindings. *Strategic Management Journal, 17*(7), 571-576.

West, M. A., & Anderson, N. R. (1996). Innovation in top management teams. *Journal of Applied Psychology, 81*(6), 680.

Westerman, G., Bonnet, D., & McAfee, A. (2014a). *Leading digital: Turning technology into business transformation* Harvard Business Press.

Westerman, G., Bonnet, D., & McAfee, A. (2014b). The nine elements of digital transformation. *MIT Sloan Management Review, 55*(3), 1-6.

Wiersema, M. F., & Bantel, K. A. (1992). Top management team demography and corporate strategic change. *The Academy of Management Journal, 35*(1), 91-121. doi:10.2307/256474

Wright, P. M., Coff, R., & Moliterno, T. P. (2014). Strategic human capital: Crossing the great divide. *Journal of Management, 40*(2), 353-370.

APPENDIX A1: List of Websites related to digital transformation, from where the preliminary list of terms were extracted

https://www.coresystems.net/blog/understanding-the-lingo-of-the-digital-transformation
https://www.vmware.com/topics/glossary/
https://digitaltransformation.net/en/glossary/
https://risnews.com/digital-transformation-top-10-terms-defining-future-retail (Aperion Glossary)
https://blogs.starcio.com/2016/07/defining-digital-transformation.html
https://medium.com/digitaladoption101/the-ultimate-dictionary-of-digital-buzzwords-2d7885de5cd0
https://www.cognizant.com/glossary/oil-gas-digital-transformation
https://www.lbbw.de/articlepage/understanding-markets/key-terms-digitalization-industry-4-0_6vip5a4gw_e.html
https://consumergoods.com/10-terms-defining-digital-transformation-consumer-goods-industry
https://www.henkel.com/company/henkelx/digital-glossary
https://csnews.com/digital-transformation-top-10-terms-defining-future-c-stores
https://www.inspirage.com/2019/04/the-ultimate-glossary-of-terms-about-digital/
https://www.essentracomponents.com/en-us/news/guides/industry-40-the-terms-you-need-to-know
www.walkme.com

Term	Frequency
artificial intelligence	11
internet of things/ IoT	11
big data/ big data analytics	10
Augmented Reality AR	8
Digital Transformation	8
cloud automation/ enablement/migration/ networking/security/services	7
Data Mining/ Data Analytics	7
machine learning	7
automation/ automation platform/ automation solutions	6
Blockchain	6

APPENDIX B1: Search parameters and sample results for the search for academic papers on Web of Science

Database: Web of Science Core Collection **ALL FIELDS:** ("digital transformation") **Refined By: DOCUMENT TYPES:** (ARTICLE OR PROCEEDINGS PAPER) AND **WEB OF SCIENCE CATEGORIES:** (COMPUTER SCIENCE INFORMATION SYSTEMS OR COMPUTER SCIENCE THEORY METHODS OR BUSINESS FINANCE OR MANAGEMENT OR COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS OR SOCIAL SCIENCES MATHEMATICAL METHODS OR BUSINESS OR SOCIAL SCIENCES INTERDISCIPLINARY OR ECONOMICS OR PSYCHOLOGY APPLIED OR COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE OR COMPUTER SCIENCE CYBERNETICS OR OPERATIONS RESEARCH MANAGEMENT SCIENCE OR BEHAVIORAL SCIENCES)

What Is Different About Digital Strategy? From Quantitative to Qualitative Change
By: Adner, Ron; Puranam, Phanish; Zhu, Feng
STRATEGY SCIENCE Volume: 4 Issue: 4 Special Issue: SI Pages: 253-261 Published: DEC 2019

How to Develop a Digital Ecosystem: a Practical Framework
By: Valdez-De-Leon, Omar
TECHNOLOGY INNOVATION MANAGEMENT REVIEW Volume: 9 Issue: 8 Pages: 54-43 Published: AUG 2019

Digital transformation, digital dividends and entrepreneurship: A quantitative analysis
By: Galindo-Martin, Miguel-Angel; Castano-Martinez, Maria-Soledad; Mendez-Picazo, Maria-Teresa
JOURNAL OF BUSINESS RESEARCH Volume: 101 Pages: 522-527 Published: AUG 2019

Demystifying AI: What Digital Transformation Leaders Can Teach You about Realistic Artificial Intelligence
By: Brock, Jurgen Kai-Uwe; von Wangenheimz, Florian
CALIFORNIA MANAGEMENT REVIEW Volume: 61 Issue: 4 Special Issue: SI Pages: 110-134 Article Number: UNSP 1536504219865226 Published: AUG 2019

Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal
By: Warner, Karl S. R.; Waeger, Maximilian
LONG RANGE PLANNING Volume: 52 Issue: 3 Pages: 326-349 Published: JUN 2019

The impacts of digital transformation on the labour market: Substitution potentials of occupations in Germany
By: Dengler, Katharina; Matthes, Britta
TECHNOLOGICAL FORECASTING AND SOCIAL CHANGE Volume: 137 Pages: 304-316 Published: DEC 2018

LOW-LOW (LL) HIGH HUMAN CAPITAL CLUSTERS IN PUBLIC ADMINISTRATION EMPLOYMENT - PREDICTOR FOR DIGITAL INFRASTRUCTURE PUBLIC INVESTMENT PRIORITY - ROMANIA CASE STUDY
By: Linearu, Cristina; Pirciog, Speranta; Grigorescu, Adriana; et al.
ENTREPRENEURSHIP AND SUSTAINABILITY ISSUES Volume: 6 Issue: 2 Pages: 753-729 Published: DEC 2018

Unpacking the Disruption Process: New Technology, Business Models, and Incumbent Adaptation
By: Cozzolino, Alessio; Verona, Gianmario; Rothaermel, Frank T.
JOURNAL OF MANAGEMENT STUDIES Volume: 55 Issue: 7 Special Issue: SI Pages: 1166-1202 Published: NOV 2018

Social Media, Social Capital, and Knowledge Sharing in Enterprise
By: Jarrahi, Mohammad Hossein
Conference: 42nd Annual IEEE-Computer-Society Computers, Software and Applications (COMPSAC) Conference - Staying Smarter in a Smartening World Location: Tokyo, JAPAN Date: JUL 23-27, 2018
Sponsor(s): IEEE Comp Soc
IT PROFESSIONAL Volume: 20 Issue: 4 Pages: 37-45 Published: JUL-AUG 2018

How AUDI AG Established Big Data Analytics in Its Digital Transformation
By: Dremel, Christian; Herterich, Matthias; Wulf, Jochen; et al.
MIS QUARTERLY EXECUTIVE Volume: 16 Issue: 2 Pages: 81-100 Published: JUN 2017

Futures of robotics. Human work in digital transformation
By: Kaivo-oja, Jari; Roth, Steffen; Westerlund, Leo
INTERNATIONAL JOURNAL OF TECHNOLOGY MANAGEMENT Volume: 73 Issue: 4 Pages: 176-205 Published: 2017

How transformational leadership facilitates e-business adoption
By: Alos-Simo, Lirios; Verdu-Jover, Antonio J.; Gomez-Gras, Jose-Maria
INDUSTRIAL MANAGEMENT & DATA SYSTEMS Volume: 117 Issue: 2 Pages: 382-397 Published: 2017

The sharing economy: Your business model's friend or foe?
By: Kathan, Wolfgang; Matzler, Kurt; Veider, Viktoria
BUSINESS HORIZONS Volume: 59 Issue: 6 Special Issue: SI Pages: 663-672 Published: NOV-DEC 2016

Does mutuality matter? Examining the bilateral nature and effects of CEO-CIO mutual understanding
By: Benlian, Alexander; Haffke, Ingmar

JOURNAL OF STRATEGIC INFORMATION SYSTEMS Volume: 25 Issue: 2 Pages: 126-104 Published: JUL 2016

Options for Formulating a Digital Transformation Strategy
By: Hess, Thomas; Matt, Christian; Benlian, Alexander; et al.
MIS QUARTERLY EXECUTIVE Volume: 15 Issue: 2 Pages: 123-139 Published: JUN 2016

APPENDIX B2: Sample terms extracted from academic literature

KEY TERM	FREQUENCY	NO. CASES
DIGITAL TECHNOLOGY/ TECHNOLOGIES	151	99
BIG DATA	106	56
INTERNET OF THINGS IOT	106	78
DIGITAL ECONOMY	92	42
ENTERPRISE ARCHITECTURE/ ARCHITECTURES EA	61	43
INFORMATION SYSTEM/ SYSTEMS	58	43
INFORMATION TECHNOLOGY/ TECHNOLOGIES	53	54
CYBER PHYSICAL SYSTEMS	36	30
REAL TIME	36	23
ARTIFICIAL INTELLIGENCE AI	34	31

APPENDIX C1: List of books from which key terms were extracted

Name Of Book	Published	Author	Goodreads.com Rating	Source Weblink
Digital Transformation Playbook: Rethink Your Business for the Digital Age	2016	David L. Rogers	avg rating 4.16 — 430 ratings	https://idoc.pub/
Leading Digital: Turning Technology into Business Transformation	2014	George Westerman	avg rating 3.81 — 547 ratings	https://www.vadira.de/hubfs/Blog%202019/02%20Februar/Leading-Digital-George-Westerman-Didier-Bonnet-And-Andrew-Mcafee.pdf
The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies	2014	Erik Brynjolfsson	avg rating 3.95 — 10,242 ratings	https://idoc.pub/
Digital Transformation at Scale: Why the Strategy Is Delivery	2018	Andrew Greenway	avg rating 4.25 — 63 ratings	https://sites.google.com/site/themermaidandmrshancockpdf3/book-pdf-digital-transformation-at-scale-why-the-strategy-is-delivery-perspectives-by-andrew-greenway-77514aa213
Managing Digital transformation	2018	Per Andersson et al	N/A	https://www.hhs.se/contentassets/a3083bb76c384052b3f3f4c82236e38f/managing-digital-transformation-med-omslag.pdf
Digital Transformation for Dummies	2017	Zeus Kerravala, Lawrence Miller	N/A	https://moodle.ufsc.br/pluginfile.php/2788927/mod_resource/content/1/2017%20-%20Digital%20Transformation%20for%20Dummies%20-%20Kerravala%20and%20Miller.pdf
Digital Transformation with BPM	2019	Nathaniel Palmer et al	N/A	http://documents.bpm.com/digitaltransformation/DigitalTransformationDigitalEdition.pdf

APPENDIX C2: Sample key terms extracted from books on digital transformation

Word/ Term	Count
digital technology	222
customer experience	177
big data	151
social media	150
business process	150
real time	102
digital vision	80
digital age	74
platform business	72
digital business	70

APPENDIX D: Template sent to subject matter experts

Your Opinion on Digital Transformation related terms
My name is Tushar R. Shah, and I am requesting your participation as a subject matter expert (SME) in my dissertation study being conducted under the auspices of UT Arlington. My dissertation study is titled, "Human Capital and Social Capital of Upper Echelons: Their influence on strategic change decisions involving digital transformation". This dissertation studies organizational change in the context of digital transformation. Specifically, I examine the role of organization's upper echelons (i.e. top management teams) in predicting digital transformation. I seek your help in identifying the domain of "digital transformation".
Your opinion will be explicitly acknowledged in my dissertation document. However, for that to happen, I am required to seek your consent because the Institutional Review Board (IRB) of my university (University of Texas at Arlington) asks me to do so. The Institutional Review Board is a body which controls how human subjects will be treated during the process of seeking their opinions on certain research questions. Some of the statements that follow may seem strange, repetitive, overwhelming and/or out-of-place, but those are a part of the standard IRB protocol.
I try to determine the domain of "digital transformation" by seeking your specific opinion on a glossary of terms/ phrases. As you are among the identified subject matter experts on this topic, your opinion is very valuable. By participating in this study, I will benefit immensely with your expertise on this topic. By participating in this study, you may provide an academic contribution in terms of defining the domain. Yet, responding to this email is voluntary.
With you agreeing to do this exercise, you acknowledge and accept the following:
1. Requirement of this informed consent acceptance is a part of standard protocol for seeking opinions from subject matter experts
2. There are no known risks for participation or adverse consequences for not participating in this study
3. Your name will be explicitly acknowledged in the dissertation document
4. Your response combined with responses from other participants will form the basis of my claim that subject matter experts have validated the different dimensions of what "digital transformation" means
This study has been reviewed and approved by the University of Texas at Arlington Institutional Review Board. It is under the supervision of Tushar Shah. Please feel free to contact him on tushar.shah@mavs.uta.edu or on mobile number 682-330-0089. If you want to contact his faculty advisor, you can email your communication to Dr. George Benson on benson@uta.edu
The Main Question of my study to which I seek your opinion:

On a scale of 0 to 5, please rate the following terms on their association
with Digital Transformation, where 0 = Not Associated and 5 = Highly Associated
This exercise will take about 25-30 minutes of your time.

Basic Information About Respondent

1 Total Years of Experience:

2 Academic Scholar OR Corporate Professional (OR Both):

3 Industry/ Industrial Sector Currently associated with:

4 Your Role OR Type of digital transformation-related activity you are associated with (in a single sentence) : E.g.: I am the CEO of a company which provides data analytics solutions to mid-size companies

APPENDIX D Continued

0 = Not Associated and 5 = Highly Associated

In case of any ambiguity, you can rate the term/ phrase based on whatever context comes to your mind after you see the word. If you feel some word/s is/are left out, you may add them into columns C, D, and E on row 116. If you want to add more than 3 words, please add them in rows 117 onwards

Keywords/ Phrases1	Rating1	Keywords/ Phrases2	Rating2	Keywords/ Phrases3	Rating3	Keywords/ Phrases4	Rating4
3D Planogram/planogram automation		Data Science		Endless Aisle		Network Analytics/ Analysis	
3D Space Management		Data Scientists		Enterprise application services		Network Virtualization	
Adaptive Case Management		Decision Making Process		Enterprise Architecture/ Architectures EA		Offline touch point	
Adaptive Integrated EA Framework		Decision Support System		Enterprise Cloud		omni-channel retail/ omni-channel retail identity management	
Additive Manufacturing		Deep Learning		Enterprise collaboration		one-click check-out	
advanced analytics		delivery drone		Enterprise Engineering		Online marketplace/ online order management system/ on-line touchpoint	
Age Of Digital		demand forecasting		Enterprise Resource Planning ERP		Online Survey	

agile methods/development		Design Methodology Approach		Enterprise Services		online training	
agile project management		Design Process		Enterprise Transformation		Open Access	
algorithm		Design Science		Equity Crowdfunding Platforms		Open Innovation	
Algorithmic Business		Development Of Digital		ERP Systems		Open Source Software	
Amazon Retail Analytics (ARA)		DevOps		Ethereum		open source/ open data/ Open Systems Interconnection Model (OSIM)	
Amazon Web Services (AWS)		Digital Service/ Services		Facial Recognition		overall equipment effectiveness	
Anti-fraud tool		digital adoption		Financial Technologies		PAAS (Integrated Platform as a service)	
API Centric		Digital Age		Fourth Industrial Revolution		personalized marketing/ personalized shopping experience	
App/ Application Service Provider		digital alibi		Future Software		Physical And Virtual	
Application Development		Digital And Physical		gamification		Platform Based	
Application Programming Interface API		digital asset management		general digital maturity		Plug & Work	

Artificial Intelligence AI		Digital automation platform		Geo-fencing		POS system	
automated risk analysis platform (ARAP)		Digital Banking		gig economy		predictive analysis/ analytics/ modeling	
Automated System		Digital Business Model/ Strategy/ Transformation		Global Data Synchronization Network		predictive maintenance	
automation		Digital Capability/ Capabilities		hackathon		prescriptive analysis/ analytics	
automation platform		digital channel		Hadoop		Pricing optimization/ pricing algorithm	
automation solutions		digital competitiveness		High Tech		Process Innovation	
Autonomous Intelligent Vehicles		Digital Content		horizontal/ vertical integration		Process Model/ Models	
Autonomous shopping cart		Digital Context		human-machine interface (HMI)		product information management (PIM)	
Backend systems/ Backoffice applications		digital coupon/ digital wallet		Hybrid Cloud		QR code	
Barcode		Digital Culture		hyper converged technology		quantum computing	
Base Technologies		digital customer engagement		Hyper relevance		RAMI 4.0	
Beacon		Digital Data		IAAS (Infrastructure as a service)		Ransomware	
Big Data		Digital Design		Identity and access management		Real Time	

Big Data Analytics		digital destination		Image analytics/ video analytics		Realtime data (RTD)	
bitcoin		Digital Development		Impact Of Digital		Reference Models	
Blockchain		Digital Disruption		Implementation Of Digital		Resilient Run Time Environments	
Bluetooth/ BLE		digital disruptors		Indoor positioning system		Rewards mall	
Building Management System		Digital Dividends		Industrial Internet Of Things IIoT		RFID/ Radio frequency Identification	
Business Analytics		Digital Economy		Industry 4.0		robotic process automation/ RPA	
Business Digitalization		Digital Ecosystem/ Ecosystems		Industry Technologies		SaaS (Software as a Service)	
Business Ecosystem		Digital Education/ Learning		Information Age		Scaled Agile Framework (SAFe)	
Business Intelligence		Digital Engineering		Information And Communication		Scan and pay	
Business Model Innovation		Digital Enterprise		Information And Communication Technology ICT		Search Engine Marketing/ Optimization	
Business Network Based Value Creation		Digital Enterprise Architecture/ Architectures		Information Management		Self-checkout/ Self-Serve	
Business Process as a Service (BPaaS)		Digital Environment		Information Resources		Semantic and Syntactic Interoperability	
Business Process Management BPM		Digital Era		Information Security		Server Message Block	
BYOD (bring your own device)		Digital Factories		Information Society		server virtualization	

Capability Maturity		Digital Government		Information System/ Systems		Service Innovation	
card reader		Digital Information		Information Technology/ Technologies		Service Oriented Enterprise Architectures	
contactless payment		Digital Infrastructure		Innovation Diffusion		Showrooming	
chatbot		digital innovation/ disruptors		Innovation Infrastructure		Smart Cities	
Chief Analytics Officer		digital interoperability		Innovation Management		Smart City	
Chief Digital Officer CDO		Digital Leadership		Innovation Processes		Smart Devices	
Classical EE Paradigm		Digital Literacy		Innovative Development		smart factory/smart manufacturing	
cloud automation		Digital management platform		Innovative IT Projects		Smart Governance	
Cloud Based		Digital Marketing		Innovative Solutions		smart meter/ smart packaging/ smart shelf/ smart poster/ smart price tag	
Cloud Computing		Digital Maturity		Integrated hardware traceability		Smart Mobility	
cloud enablement		Digital Media		Integrated supply chain		Smart Products	
cloud migration		digital opportunities		Integrated Workplace Management System		smart retail/ smart store/ smart city/ smart village	
Cloud Mobile		Digital Organization		Intelligent Business Services And Related		Smart Service	

cloud networking		Digital Orientation		Intelligent Devices		Smart Technologies	
Cloud Platform		Digital Platform/ Platforms		Interactive screen, Interactive Voice Response		social commerce	
cloud robotics		digital practices		internet of things IoT		Social Media	
cloud security		digital process automation		Internet Technologies		social media marketing	
Cloud Services		Digital Product/ Products		Internet Users		Social Network Analysis	
cloud/ electronic POS		Digital Quality Of Life		IT Automation		social networking/ networks	
cobots		Digital Readiness		ITG Mechanisms		Social Wi-Fi	
Cognitive Computing		Digital Revolution		journey analytics/ journey map/ customer journey map		Software Architecture	
Cognitive Technologies		digital shelf edge/ electronic shelf edge/ electronic rack edge tag/ electronic shelf label		keyword stuffing		Software defined Wide Area Network (SD-WAN)	
Computer Based		Digital Skills		kinetic presenter		Software Development/ Engineering/ Systems/ Implementation	
Computer Games		Digital Society		Knowledge Based		Software Tutorial	
computerized inventory system		Digital Solutions		Knowledge Intensive		Source Code	
connected customer		Digital Space		Knowledge Management		State Of The Art	

consumer experience		digital specialist		Knowledge Management Systems		storage area network	
Content Analysis		digital store		Knowledge Networks		Streaming Analytics	
continuous replenishment process		Digital Strategy/ Strategies		Knowledge Sharing		Supervised Learning	
Control System		digital supply chain		Lean Management		System Design	
convenience technology		Digital Technology/ Technologies		Lean Office		Task Automation	
conversion rate optimization		Digital Tools		Lean production		Technology /Technological Development	
CPV (cost per view)		digital traceability		Level Of Digitalization		Technology adoption	
Crowd Modeling		Digital Transformation		Li-Fi (Light fidelity)		Technology Architecture	
crowdsourcing		digital twin		lights out environment/ lights out manufacturing		Technology Driven	
crowdfunding		digital user journey		Linked Open Data		Tokenization	
CTR (click through rate)		digital value-added		Liquid application		Traffic Counter	
Current Digital		digital vision		live label		Training software	
customer analytics		Digital Workplace		LTE (Long term evolution)		UAV (Unmanned Aerial Vehicle)/ Drone	
Customer Data Management		Digital World		Machine Learning		unified commerce platform/ Unified retail	
customer engagement analytics		Digitally Enabled Innovation		machine to machine		User Centered	

customer helping robots		Digitization On Risk Management		manufacturing execution systems		User Experience/User Interface Design (UX/UI)/ User Onboarding	
customer intelligence		digitization/ digitalization		Mapping Study		Utility Analytics	
Customer Relationship Management CRM		Digitized Industrial Products		Marketing Automation		Values Models	
Cyber Attacks		DIPSE (digital interactive predictive sensory edge)		Maturity Model		Video Content	
Cyber Infrastructures		Disruptive Innovation/ Technologies		merchant portal		Virtual And Augmented Reality VR AR	
Cyber Physical Systems		Distributed Information Systems		micro-blog		virtual assistant	
Cyber security		distributed order management		micro-services		virtual machine	
cyber-physical production systems		E-commerce/ E-tailing/ M-commerce		middleware		virtual network	
dark data		Edge Computing		mixed reality		virtual omnichannel inventory management	
data		EE Paradigm		Mobile Apps		Virtual World	
Data Analysis		E-learning		Mobile Devices		Virtualization Of Consulting Services	

Data Analytics		Electronic Article Surveillance		Mobile Health		visible light communication (VLC)/ visual recognition system	
Data And Information		Electronic Data Interchange EDI		mobile POS, mobile shopping, mobile wallet		voice commerce	
data architecture		Electronic Governance		Mobile Technologies		voice controlled AI/ conversational AI	
data center		Electronic product code (EPC), Information Services (EPCIS)		Mobile Telephone		voice search/ voice activated assistant	
Data Centric		Embedded systems		Mobility Systems		walkthroughs	
Data Driven		Emerging Technologies		Model Driven		warehouse management system (WMS)	
Data Governance		Empirical Research		Model-driven workflow		wearable technology	
Data Lake		EMV Technology		Modern Information Technologies		Web Applications	
Data Management		Enabling Technologies		multi-factor authentication		webrooming	
Data Migration		Encryption		Natural Language Processing		website navigation	
Data Mining		End User		near field communication (NFC)		Work Design	
Data Processing							

APPENDIX E: Final list of keywords after SME validation and academic scholars vetting

Keywords/ Phrases1	Keywords/ Phrases2	Keywords/ Phrases3	Keywords/ Phrases4
3D Planogram/planogram automation	Data Science	Encryption	Open Source Software
3D Space Management	Data Scientists	Enterprise Architecture/ Architectures EA	open source/ open data/ Open Systems Interconnection Model (OSIM)
Adaptive Integrated EA Framework	Decision Making Process	Enterprise Cloud	PAAS (Integrated Platform as a service)
Additive Manufacturing/ 3D Printing	Decision Support System	E-learning	personalized marketing/ personalized shopping experience
advanced analytics	Deep Learning	Electronic Article Surveillance	Platform Based
agile methods/development	delivery drone	Electronic Governance	Plug & Work
Amazon Retail Analytics (ARA)	demand forecasting	Electronic product code (EPC), Information Services (EPCIS)	predictive analysis/ analytics/ modeling
Amazon Web Services (AWS)	Design Methodology Approach	Embedded systems	predictive maintenance
Anti-fraud tool	Design Science	Emerging Technologies	prescriptive analysis/ analytics
API Centric	DevOps	Enterprise collaboration	Pricing optimization/ pricing algorithm
App/ Application Service Provider	Digital Service/ Services	Enterprise Resource Planning ERP, ERP Systems	Process Innovation
Application Programming Interface API	digital adoption	Enterprise Transformation	product information management (PIM)
Artificial Intelligence AI	Digital Age, Digital Era	Equity Crowdfunding Platforms	QR code
automated risk analysis platform (ARAP)	digital alibi	Ethereum	quantum computing
Automated System	Digital And Physical	Facial Recognition	Ransomware
automation	digital asset management	Fourth Industrial Revolution	Real Time

automation platform	Digital automation platform	gamification	Realtime data (RTD)
automation solutions	Digital Banking	general digital maturity	Reference Models
Autonomous Intelligent Vehicles	Digital Business Model/ Strategy/ Transformation	Geo-fencing	Resilient Run Time Environments
Autonomous shopping cart	Digital Capability/ Capabilities	gig economy	RFID/ Radio frequency Identification
Beacon	digital channel	Global Data Synchronization Network	robotic process automation/ RPA
Big Data	digital competitiveness	hackathon	SaaS (Software as a Service)
Big Data Analytics	Digital Content	Hadoop	Scaled Agile Framework (SAFe)
bitcoin	digital coupon/ digital wallet	High Tech	Scan and pay
Blockchain	Digital Culture, Digital Society	horizontal/ vertical integration	Search Engine Marketing/ Optimization
Bluetooth/ BLE	digital customer engagement	human-machine interface (HMI)	Self-checkout/ Self-Serve
Building Management System	Digital Data	Hybrid Cloud	Semantic and Syntactic Interoperability
Business Analytics	Digital Design	hyper converged technology	Server Message Block
Business Digitalization	digital destination	Hyper relevance	server virtualization
Business Intelligence	Digital Development/ Digital Engineering	IAAS (Infrastructure as a service)	Service Innovation
Business Model Innovation	Digital Disruption, digital disruptors	Identity and access management	Service Oriented Enterprise Architectures
Business Network Based Value Creation	Digital Dividends	Image analytics/ video analytics	Smart Devices
Business Process as a Service (BPaaS)	Digital Economy	Indoor positioning system	smart factory/smart manufacturing
Business Process Management BPM	Digital Ecosystem/ Ecosystems	Industrial Internet Of Things IIoT	Smart Governance
BYOD (bring your own device)	Digital Education/ Learning	Industry 4.0	smart meter/ smart packaging/ smart shelf/ smart poster/ smart price tag

Capability Maturity	Digital Enterprise Architecture/ Architectures, Digital Infrastructure	Industry Technologies	Smart Mobility
card reader	Digital Environment, Digital World, Digital Context	Information Security	Smart Products
contactless payment	Digital Factories	Innovation Infrastructure, Innovation Processes, Innovation Diffusion	smart retail/ smart store/ smart city/ smart village
chatbot	Digital Government	Innovative IT Projects	Smart Service
Chief Analytics Officer	Digital Information	Integrated supply chain	Smart Technologies
Chief Digital Officer CDO	digital innovation, Disruptive Innovation/ Technologies, Digitally Enabled Innovation	Integrated Workplace Management System	social commerce
cloud automation	digital interoperability	Intelligent Business Services And Related	Social Media
Cloud Based	Digital Leadership	Intelligent Devices	social media marketing
Cloud Computing, cloud enablement, Cloud Platform, cloud security, Cloud Services	Digital Literacy	Interactive screen, Interactive Voice Response	Social Network Analysis
cloud/ electronic POS	Digital management platform	internet of things IoT	social networking/ networks
cloud migration, Cloud Mobile	Digital Marketing	Internet Technologies	Social Wi-Fi
cloud networking, cloud robotics	Digital Maturity	IT Automation	Software defined Wide Area Network (SD-WAN)
cobots	Digital Media	ITG Mechanisms	Software Development/ Engineering/ Systems/ Implementation
Cognitive Computing	Digital Orientation	journey analytics/ journey map/	storage area network

		customer journey map	
Cognitive Technologies	Digital Platform/ Platforms	keyword stuffing	Streaming Analytics
connected customer	digital practices	Knowledge Intensive	Supervised Learning
consumer experience	digital process automation	Knowledge Networks	Task Automation
Content Analysis	Digital Product/ Products, Digital Solutions	Knowledge Sharing	Technology /Technological Development
conversion rate optimization	Digital Quality Of Life	Level Of Digitalization	Technology adoption
CPV (cost per view)	Digital Readiness	Li-Fi (Light fidelity)	Technology Architecture
Crowd Modeling	Digital Revolution	lights out environment/ lights out manufacturing	Technology Driven
crowdsourcing, crowdfunding	digital shelf edge/ electronic shelf edge/ electronic rack edge tag/ electronic shelf label	Linked Open Data	Tokenization
CTR (click through rate)	Digital Skills	LTE (Long term evolution)	UAV (Unmanned Aerial Vehicle)/ Drone
Current Digital	Digital Space	Machine Learning	unified commerce platform/ Unified retail
customer analytics, customer engagement analytics	digital specialist	machine to machine	User Centered
Customer Data Management	digital store	Marketing Automation	User Experience/User Interface Design (UX/UI)/ User Onboarding
customer helping robots	Digital Strategy/ Strategies	Maturity Model	Utility Analytics
customer intelligence	digital supply chain	micro-blog	Virtual And Augmented Reality VR AR
Customer Relationship Management CRM	Digital Technology/ Technologies	micro-services	virtual assistant
Cyber Attacks	Digital Tools	middleware	virtual machine
Cyber Infrastructures	digital traceability	mixed reality	virtual network

Cyber Physical Systems	Digital Transformation	Mobile Apps	virtual omnichannel inventory management
Cyber security	digital twin	Mobile Devices	Virtual World
cyber-physical production systems	digital user journey	Mobile Health	Virtualization Of Consulting Services
Data Analysis	digital value-added	Mobile Technologies, mobile POS, mobile shopping, mobile wallet	visible light communication (VLC)/ visual recognition system
Data Analytics	digital vision	Modern Information Technologies	voice commerce
Data And Information	Digitization On Risk Management	multi-factor authentication	voice controlled AI/ conversational AI
data architecture	digitization/ digitalization	Natural Language Processing	voice search/ voice activated assistant
Data Centric	Digitized Industrial Products	near field communication (NFC)	Video Content
Data Driven	DIPSE (digital interactive predictive sensory edge)	Network Analytics/ Analysis	warehouse management system (WMS)
Data Governance	Distributed Information Systems	Network Virtualization	wearable technology
Data Lake	distributed order management	omni-channel retail/ omni-channel retail identity management	Web Applications
Data Management	E-commerce/ E-tailing/ M-commerce	one-click check-out	webrooming
Data Migration	Edge Computing	Online marketplace/ online order management system/ on-line touchpoint	website navigation
Data Mining and Data Processing	EE Paradigm	Open Innovation	Work Design

APPENDIX F: Python code to download annual reports from annualreports.com website

(Written by Yugesha Sapte)

```
import pandas as pd
import config as cfg
import selenium
from selenium import webdriver
import time
from bs4 import BeautifulSoup
import requests
from selenium.webdriver.common.by import By
import os
from xlwt import Workbook
import xlrd
from openpyxl import load_workbook
from selenium.common.exceptions import NoSuchElementException

nyseList = []
index_company = 0
company = ""
ipo_prospectus = ""
file_name = cfg.company_list_filename
sheet_name = cfg.sheet_name
col_name = cfg.col_name
driver = webdriver.Chrome(
    executable_path=os.path.join(os.getcwd(), cfg.chrome_driver))

def checkIfCompanyMatches(company):
    try:

        response = requests.get(driver.current_url)
        ipo_prospectus = BeautifulSoup(response.content)
        flist = []
        main_content = ipo_prospectus.find('table')
        if(main_content != None):
            rows = main_content.findAll('tr')
            for tr in rows:
                td = tr.find_all('td')
                row = [tr.text.strip() for tr in td]
                if(len(row) != 0):

                    # if(( nyseList['company'] == row[0]).any()):
                    if (company.lower() in row[0].lower()):
                        return True, row[0]
                    else:
```

```
                company_update("Not in NASDAQ list!")
                return False, row[0]
        else:
            company_update("No company name found!")
            return False, False

    except Exception as e:
        company_update("No company name found!")
        print("Exception:", company, " name not found!")
        pass

def set_company_details(tickerId):
    try:
        company_name = driver.find_element_by_name(cfg.company)
        company_name.send_keys(tickerId)
        time.sleep(cfg.sleep_time)
        search_button = driver.find_element_by_xpath(
            "/html/body/div/section[1]/div/form/fieldset/div[1]/input[2]")
        search_button.click()
    except:
        pass

def company_update(status):

    df = pd.DataFrame([status])
    writer = pd.ExcelWriter(file_name, engine='openpyxl')
    writer.book = load_workbook(file_name)
    writer.sheets = {ws.title: ws for ws in writer.book.worksheets}
    df.to_excel(writer, sheet_name=sheet_name, startrow=index_company +
            1, startcol=2, index=False, header=False)
    writer.save()

def read_company_list():
    df = pd.read_excel(file_name, sheet_name=sheet_name)
    return df

def navigate_to_company_page(driver, company):
    pdf_list = {}
    try:

        link = driver.find_element_by_link_text(company)
        link.click()

        divv = driver.find_element_by_class_name('content-archive')

        content_archive = divv.find_elements_by_class_name('links')
```

```
        for item in content_archive:

            link = item.find_element_by_css_selector('a')
            text = link.get_attribute('href').split('/')

            pdf_list[text[len(text)-1]] = link.get_attribute('href')

    except Exception as e:
        company_update("Issue in listing documents")

    finally:

        return pdf_list

def save_archieved_pdfs(urls, company_name):
    try:
        company_name = company_name.replace(' ', '')
        company_name = company_name.replace(',', '')
        path = "reports/"+company_name
        os.mkdir(path)

        for url in urls:
            image_url = urls[url]
            r = requests.get(image_url)

            with open(os.path.join(path, url), 'wb') as f:
                f.write(r.content)

        company_update("Success!")

    except Exception as e:
        company_update("Issue in saving documents")

def main():

    # unique name list
    # df = pd.read_excel(file_name, sheet_name=sheet_name)
    # df = df.drop_duplicates('Company Name')
    # writer = pd.ExcelWriter('finallist1.xlsx', engine='openpyxl')
    # writer.book = load_workbook('finallist1.xlsx')
    # df.to_excel(writer, index=False)
    # writer.save()
    global nyseList
    nyseList = pd.read_excel("NYSEList.xlsx", sheet_name="Sheet1")
    i = 0
    df = pd.read_excel(file_name, sheet_name=sheet_name)
```

```
    for index, row in df.iterrows():
        if(pd.isnull(row["Status"])):
            flag = True
            global index_company
            index_company = index
            company = row["Company Name"]
            # row['Ticker Symbol']
            if(i == 150):
                break
            driver.get(cfg.web_url)
            time.sleep(5)
            set_company_details(company)
            time.sleep(5)
            flag, name = checkIfCompanyMatches(company)
            if(flag):
                docs = navigate_to_company_page(driver, name)
                # spf.save_archieved_pdfs(docs, name)
                save_archieved_pdfs(docs, name)

            time.sleep(5)
            driver.close
            i += 1

main()
```

config.py

```
company_list_filename = "finallist.xlsx"
sheet_name = "WRDS"
# "544_Companies-2009-2015"
# "1483_Companies_2006-2018"
col_name  = "Company Name"
chrome_driver = "chromedriver.exe"
web_url  = 'http://annualreports.com/'
company = 'search'
search_button_1 ='search_button_1'
type = 'type'
doc_type = '424b'
period = 'prior_to'
period_year = '2018'
text = 'Documents'
link_address = "href"
sleep_time = 5
text1 = '.htm'
```

APPENDIX G: Python code to score annual reports on lexicon for digital transformation (Written by Sanjana Suresh)


```
{
 "cells": [
  {
   "cell_type": "code",
   "execution_count": 1,
   "metadata": {},
   "outputs": [],
   "source": [
    "import pandas as pd"
   ]
  },
  {
   "cell_type": "code",
   "execution_count": 2,
   "metadata": {},
   "outputs": [
    {
     "data": {
      "text/html": [
       "<div>\n",
       "<style scoped>\n",
       "    .dataframe tbody tr th:only-of-type {\n",
       "        vertical-align: middle;\n",
       "    }\n",
       "\n",
       "    .dataframe tbody tr th {\n",
       "        vertical-align: top;\n",
       "    }\n",
       "\n",
       "    .dataframe thead th {\n",
       "        text-align: right;\n",
       "    }\n",
       "</style>\n",
       "<table border=\"1\" class=\"dataframe\">\n",
       "  <thead>\n",
       "    <tr style=\"text-align: right;\">\n",
       "      <th></th>\n",
       "      <th>Keywords/ Phrases1</th>\n",
       "      <th>Rating1</th>\n",
       "    </tr>\n",
       "  </thead>\n",
       "  <tbody>\n",
```

```
"   <tr>\n",
"     <th>0</th>\n",
"     <td>3D Planogram</td>\n",
"     <td>3.714286</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>1</th>\n",
"     <td>planogram automation</td>\n",
"     <td>3.714286</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>2</th>\n",
"     <td>3D Space Management</td>\n",
"     <td>3.571429</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>3</th>\n",
"     <td>Adaptive Integrated EA Framework</td>\n",
"     <td>2.142857</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>4</th>\n",
"     <td>Additive Manufacturing</td>\n",
"     <td>1.714286</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>...</th>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>431</th>\n",
"     <td>wearable technology</td>\n",
"     <td>3.428571</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>432</th>\n",
"     <td>Web Applications</td>\n",
"     <td>3.000000</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>433</th>\n",
"     <td>webrooming</td>\n",
"     <td>2.571429</td>\n",
"   </tr>\n",
"   <tr>\n",
```

```
      "      <th>434</th>\n",
      "      <td>website navigation</td>\n",
      "      <td>2.571429</td>\n",
      "    </tr>\n",
      "    <tr>\n",
      "      <th>435</th>\n",
      "      <td>Work Design</td>\n",
      "      <td>1.714286</td>\n",
      "    </tr>\n",
      "  </tbody>\n",
      "</table>\n",
      "<p>436 rows Ã— 2 columns</p>\n",
      "</div>"
     ],
     "text/plain": [
      "                 Keywords/ Phrases1    Rating1\n",
      "0                       3D Planogram  3.714286\n",
      "1                 planogram automation  3.714286\n",
      "2                  3D Space Management  3.571429\n",
      "3    Adaptive Integrated EA Framework  2.142857\n",
      "4              Additive Manufacturing  1.714286\n",
      "..                                ...       ...\n",
      "431              wearable technology  3.428571\n",
      "432                Web Applications  3.000000\n",
      "433                     webrooming  2.571429\n",
      "434               website navigation  2.571429\n",
      "435                     Work Design  1.714286\n",
      "\n",
      "[436 rows x 2 columns]"
     ]
    },
    "execution_count": 2,
    "metadata": {},
    "output_type": "execute_result"
   }
  ],
  "source": [
   "keywords_file = pd.read_excel('E:\\\\GTA work\\\\Tushar\\\\Report
Reading\\\\FinalKeyWordListWithWeightages.xlsx',
sheet_name='TruncFinalValidatedVerticalForm',skiprows=1)\n",
   "#keywords_file.drop(columns='Unnamed: 0', inplace=True)\n",
   "keywords_file"
  ]
 },
 {
  "cell_type": "code",
```

```
 "execution_count": 3,
 "metadata": {},
 "outputs": [],
 "source": [
  "keywords_list = keywords_file['Keywords/ Phrases1'].to_list()"
 ]
},
{
 "cell_type": "code",
 "execution_count": 4,
 "metadata": {},
 "outputs": [],
 "source": [
  "import string\n",
  "\n",
  "p = string.punctuation\n",
  "d = string.digits\n",
  "\n",
  "table = str.maketrans(p, len(p)* ' ')"
 ]
},
{
 "cell_type": "code",
 "execution_count": 5,
 "metadata": {},
 "outputs": [],
 "source": [
  "def keyword_clean(keyword_list):\n",
  "    keyword_clean_list = []\n",
  "    for word in keyword_list:\n",
  "        word_edit = word.translate(table)\n",
  "        word_edit1 = word_edit.lower()\n",
  "        #print(word_edit2.strip())\n",
  "        keyword_clean_list.append(word_edit1.strip())\n",
  "    return(keyword_clean_list)"
 ]
},
{
 "cell_type": "code",
 "execution_count": 6,
 "metadata": {},
 "outputs": [],
 "source": [
  "keywords_clean = keyword_clean(keywords_list)\n",
  "keywords_clean_pd = pd.Series(keywords_clean)\n",
  "keywords_file['Keywords Modified'] = keywords_clean_pd.values\n"
```

```
]
},
{
 "cell_type": "code",
 "execution_count": 7,
 "metadata": {},
 "outputs": [
  {
   "data": {
    "text/html": [
     "<div>\n",
     "<style scoped>\n",
     "    .dataframe tbody tr th:only-of-type {\n",
     "        vertical-align: middle;\n",
     "    }\n",
     "\n",
     "    .dataframe tbody tr th {\n",
     "        vertical-align: top;\n",
     "    }\n",
     "\n",
     "    .dataframe thead th {\n",
     "        text-align: right;\n",
     "    }\n",
     "</style>\n",
     "<table border=\"1\" class=\"dataframe\">\n",
     "  <thead>\n",
     "    <tr style=\"text-align: right;\">\n",
     "      <th></th>\n",
     "      <th>Keywords/ Phrases1</th>\n",
     "      <th>Rating1</th>\n",
     "      <th>Keywords Modified</th>\n",
     "    </tr>\n",
     "  </thead>\n",
     "  <tbody>\n",
     "    <tr>\n",
     "      <th>0</th>\n",
     "      <td>3D Planogram</td>\n",
     "      <td>3.714286</td>\n",
     "      <td>3d planogram</td>\n",
     "    </tr>\n",
     "    <tr>\n",
     "      <th>1</th>\n",
     "      <td>planogram automation</td>\n",
     "      <td>3.714286</td>\n",
     "      <td>planogram automation</td>\n",
     "    </tr>\n",
```

```
"   <tr>\n",
"     <th>2</th>\n",
"     <td>3D Space Management</td>\n",
"     <td>3.571429</td>\n",
"     <td>3d space management</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>3</th>\n",
"     <td>Adaptive Integrated EA Framework</td>\n",
"     <td>2.142857</td>\n",
"     <td>adaptive integrated ea framework</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>4</th>\n",
"     <td>Additive Manufacturing</td>\n",
"     <td>1.714286</td>\n",
"     <td>additive manufacturing</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>...</th>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>431</th>\n",
"     <td>wearable technology</td>\n",
"     <td>3.428571</td>\n",
"     <td>wearable technology</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>432</th>\n",
"     <td>Web Applications</td>\n",
"     <td>3.000000</td>\n",
"     <td>web applications</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>433</th>\n",
"     <td>webrooming</td>\n",
"     <td>2.571429</td>\n",
"     <td>webrooming</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>434</th>\n",
"     <td>website navigation</td>\n",
"     <td>2.571429</td>\n",
```

```
 "      <td>website navigation</td>\n",
 "    </tr>\n",
 "    <tr>\n",
 "      <th>435</th>\n",
 "      <td>Work Design</td>\n",
 "      <td>1.714286</td>\n",
 "      <td>work design</td>\n",
 "    </tr>\n",
 "  </tbody>\n",
 "</table>\n",
 "<p>436 rows Ã— 3 columns</p>\n",
 "</div>"
],
"text/plain": [
 "                Keywords/ Phrases1   Rating1  \\\n",
 "0                     3D Planogram  3.714286   \n",
 "1               planogram automation  3.714286   \n",
 "2                3D Space Management  3.571429   \n",
 "3    Adaptive Integrated EA Framework  2.142857   \n",
 "4              Additive Manufacturing  1.714286   \n",
 "..                                ...       ...   \n",
 "431              wearable technology  3.428571   \n",
 "432                Web Applications  3.000000   \n",
 "433                     webrooming  2.571429   \n",
 "434               website navigation  2.571429   \n",
 "435                    Work Design  1.714286   \n",
 "\n",
 "                 Keywords Modified  \n",
 "0                     3d planogram  \n",
 "1               planogram automation  \n",
 "2                3d space management  \n",
 "3    adaptive integrated ea framework  \n",
 "4              additive manufacturing  \n",
 "..                                ...  \n",
 "431              wearable technology  \n",
 "432                web applications  \n",
 "433                     webrooming  \n",
 "434               website navigation  \n",
 "435                    work design  \n",
 "\n",
 "[436 rows x 3 columns]"
]
},
"execution_count": 7,
"metadata": {},
"output_type": "execute_result"
```

```
  }
 ],
 "source": [
  "keywords_file"
 ]
},
{
 "cell_type": "code",
 "execution_count": 8,
 "metadata": {},
 "outputs": [
  {
   "data": {
    "text/plain": [
     "{'3d planogram': 3.7142857142857144,\n",
     " 'planogram automation': 3.7142857142857144,\n",
     " '3d space management': 3.5714285714285716,\n",
     " 'adaptive integrated ea framework': 2.142857142857143,\n",
     " 'additive manufacturing': 1.7142857142857142,\n",
     " '3d printing': 1.7142857142857142,\n",
     " 'advanced analytics': 4.0,\n",
     " 'agile methods': 2.0,\n",
     " 'agile development': 2.0,\n",
     " 'amazon retail analytics': 3.857142857142857,\n",
     " 'amazon web services': 4.0,\n",
     " 'anti fraud tool': 3.4285714285714284,\n",
     " 'api centric': 3.2857142857142856,\n",
     " 'application service provider': 2.857142857142857,\n",
     " 'apps': 2.857142857142857,\n",
     " 'application programming interface': 3.5714285714285716,\n",
     " 'api': 3.5714285714285716,\n",
     " 'artificial intelligence': 4.714285714285714,\n",
     " 'ai based': 4.714285714285714,\n",
     " 'automated risk analysis platform': 3.857142857142857,\n",
     " 'arap': 3.857142857142857,\n",
     " 'automated system': 4.0,\n",
     " 'automation': 3.7142857142857144,\n",
     " 'automation platform': 4.142857142857143,\n",
     " 'automation solutions': 4.166666666666667,\n",
     " 'autonomous intelligent vehicles': 4.285714285714286,\n",
     " 'autonomous shopping cart': 4.0,\n",
     " 'beacon': 2.142857142857143,\n",
     " 'big data': 4.285714285714286,\n",
     " 'big data analytics': 4.428571428571429,\n",
     " 'bitcoin': 4.285714285714286,\n",
     " 'blockchain': 4.571428571428571,\n",
```

" 'bluetooth': 3.4285714285714284,\n",
" 'ble': 3.4285714285714284,\n",
" 'building management system': 2.142857142857143,\n",
" 'business analytics': 3.5714285714285716,\n",
" 'business digitalization': 3.4285714285714284,\n",
" 'business intelligence': 3.4285714285714284,\n",
" 'business model innovation': 3.4285714285714284,\n",
" 'business network based value creation': 2.2857142857142856,\n",
" 'business process as a service': 3.5714285714285716,\n",
" 'bpaas': 3.5714285714285716,\n",
" 'business process management': 3.142857142857143,\n",
" 'bpm': 3.142857142857143,\n",
" 'byod': 3.142857142857143,\n",
" 'capability maturity': 2.4285714285714284,\n",
" 'card reader': 2.4285714285714284,\n",
" 'chatbot': 4.285714285714286,\n",
" 'chief analytics officer': 3.7142857142857144,\n",
" 'chief digital officer': 3.857142857142857,\n",
" 'cloud automation': 4.142857142857143,\n",
" 'cloud based': 3.7142857142857144,\n",
" 'cloud computing': 4.428571428571429,\n",
" 'cloud enablement': 4.428571428571429,\n",
" 'cloud platform': 4.428571428571429,\n",
" 'cloud security': 4.428571428571429,\n",
" 'cloud services': 4.428571428571429,\n",
" 'cloud migration': 4.285714285714286,\n",
" 'cloud mobile': 4.285714285714286,\n",
" 'cloud networking': 4.142857142857143,\n",
" 'cloud robotics': 4.142857142857143,\n",
" 'cloud pos': 3.5714285714285716,\n",
" 'electronic pos': 3.5714285714285716,\n",
" 'cobots': 3.2857142857142856,\n",
" 'cognitive computing': 4.142857142857143,\n",
" 'cognitive technologies': 4.166666666666667,\n",
" 'connected customer': 3.0,\n",
" 'consumer experience': 3.4285714285714284,\n",
" 'contactless payment': 3.7142857142857144,\n",
" 'content analysis': 2.7142857142857144,\n",
" 'conversion rate optimization': 2.142857142857143,\n",
" 'cpv': 2.5714285714285716,\n",
" 'cost per view': 2.5714285714285716,\n",
" 'crowd modeling': 3.142857142857143,\n",
" 'crowdsourcing': 3.4285714285714284,\n",
" 'crowd sourcing': 3.4285714285714284,\n",
" 'crowd funding': 3.4285714285714284,\n",
" 'crowdfunding': 3.4285714285714284,\n",

" 'ctr': 2.857142857142857,\n",
" 'click through rate': 2.857142857142857,\n",
" 'current digital': 3.142857142857143,\n",
" 'customer analytics': 4.0,\n",
" 'customer engagement analytics': 4.0,\n",
" 'customer data management': 3.4285714285714284,\n",
" 'customer helping robots': 3.5714285714285716,\n",
" 'customer intelligence': 3.857142857142857,\n",
" 'customer relationship management': 3.2857142857142856,\n",
" 'crm': 3.2857142857142856,\n",
" 'cyber attacks': 3.7142857142857144,\n",
" 'cyber infrastructures': 3.5714285714285716,\n",
" 'cyber physical systems': 3.4285714285714284,\n",
" 'cyber security': 4.0,\n",
" 'cyber physical production systems': 3.5714285714285716,\n",
" 'data analysis': 3.857142857142857,\n",
" 'data analytics': 4.0,\n",
" 'data and information': 3.5714285714285716,\n",
" 'data architecture': 3.5714285714285716,\n",
" 'data centric': 3.5714285714285716,\n",
" 'data driven': 3.7142857142857144,\n",
" 'data governance': 3.5714285714285716,\n",
" 'data lake': 3.0,\n",
" 'data management': 3.7142857142857144,\n",
" 'data migration': 3.142857142857143,\n",
" 'data mining and data processing': 3.4285714285714284,\n",
" 'data science': 4.142857142857143,\n",
" 'data scientists': 4.0,\n",
" 'decision making process': 2.857142857142857,\n",
" 'decision support system': 3.142857142857143,\n",
" 'deep learning': 4.142857142857143,\n",
" 'delivery drone': 3.857142857142857,\n",
" 'demand forecasting': 3.4285714285714284,\n",
" 'design methodology approach': 3.0,\n",
" 'design science': 2.142857142857143,\n",
" 'devops': 3.7142857142857144,\n",
" 'digital service': 4.0,\n",
" 'digital adoption': 4.0,\n",
" 'digital age': 3.7142857142857144,\n",
" 'digital era': 3.7142857142857144,\n",
" 'digital alibi': 2.857142857142857,\n",
" 'digital and physical': 2.7142857142857144,\n",
" 'digital asset management': 3.142857142857143,\n",
" 'digital automation platform': 3.2857142857142856,\n",
" 'digital banking': 4.428571428571429,\n",
" 'digital business model': 4.571428571428571,\n",

" 'digital business strategy': 4.571428571428571,\n",
" 'digital business transformation': 4.571428571428571,\n",
" 'digital capability': 4.142857142857143,\n",
" 'digital capabilities': 4.142857142857143,\n",
" 'digital channel': 4.142857142857143,\n",
" 'digital competitiveness': 4.0,\n",
" 'digital content': 4.0,\n",
" 'digital coupon': 3.857142857142857,\n",
" 'digital wallet': 3.857142857142857,\n",
" 'digital culture': 3.7142857142857144,\n",
" 'digital society': 3.7142857142857144,\n",
" 'digital customer engagement': 4.0,\n",
" 'digital data': 3.857142857142857,\n",
" 'digital design': 4.0,\n",
" 'digital destination': 3.5714285714285716,\n",
" 'digital development': 3.4285714285714284,\n",
" 'digital engineering': 3.4285714285714284,\n",
" 'digital disruption': 4.0,\n",
" 'digital disruptors': 4.0,\n",
" 'digital dividends': 3.142857142857143,\n",
" 'digital economy': 3.857142857142857,\n",
" 'digital ecosystem': 4.0,\n",
" 'digital education': 3.857142857142857,\n",
" 'digital learning': 3.857142857142857,\n",
" 'digital enterprise architecture': 3.5714285714285716,\n",
" 'digital infrastructure': 3.5714285714285716,\n",
" 'digital environment': 3.4285714285714284,\n",
" 'digital world': 3.4285714285714284,\n",
" 'digital context': 3.4285714285714284,\n",
" 'digital factories': 3.142857142857143,\n",
" 'digital government': 3.0,\n",
" 'digital information': 3.2857142857142856,\n",
" 'digital innovation': 3.4285714285714284,\n",
" 'disruptive innovation': 3.4285714285714284,\n",
" 'disruptive technologies': 3.4285714285714284,\n",
" 'digitally enabled innovation': 3.4285714285714284,\n",
" 'digital interoperability': 3.142857142857143,\n",
" 'digital leadership': 3.142857142857143,\n",
" 'digital literacy': 3.142857142857143,\n",
" 'digital management platform': 3.0,\n",
" 'digital marketing': 4.142857142857143,\n",
" 'digital maturity': 3.857142857142857,\n",
" 'digital media': 3.7142857142857144,\n",
" 'digital orientation': 3.5714285714285716,\n",
" 'digital platform': 3.857142857142857,\n",
" 'digital practices': 3.5714285714285716,\n",

" 'digital process automation': 3.4285714285714284,\n",
" 'digital product': 3.5714285714285716,\n",
" 'digital solutions': 3.5714285714285716,\n",
" 'digital quality of life': 3.2857142857142856,\n",
" 'digital readiness': 3.5714285714285716,\n",
" 'digital revolution': 3.7142857142857144,\n",
" 'digital shelf edge': 3.4285714285714284,\n",
" 'electronic shelf edge': 3.4285714285714284,\n",
" 'electronic rack edge tag': 3.4285714285714284,\n",
" 'electronic shelf label': 3.4285714285714284,\n",
" 'digital skills': 3.5714285714285716,\n",
" 'digital space': 3.2857142857142856,\n",
" 'digital specialist': 3.2857142857142856,\n",
" 'digital store': 3.7142857142857144,\n",
" 'digital strategy': 3.7142857142857144,\n",
" 'digital strategies': 3.7142857142857144,\n",
" 'digital supply chain': 3.2857142857142856,\n",
" 'digital technology': 4.0,\n",
" 'digital technologies': 4.0,\n",
" 'digital tools': 3.7142857142857144,\n",
" 'digital traceability': 2.7142857142857144,\n",
" 'digital transformation': 4.428571428571429,\n",
" 'digital twin': 3.2857142857142856,\n",
" 'digital user journey': 3.7142857142857144,\n",
" 'digital value added': 3.7142857142857144,\n",
" 'digital vision': 3.7142857142857144,\n",
" 'digitization on risk management': 3.4285714285714284,\n",
" 'digitization': 3.5714285714285716,\n",
" 'digitalization': 3.5714285714285716,\n",
" 'digitized industrial products': 2.857142857142857,\n",
" 'dipse': 3.4285714285714284,\n",
" 'digital interactive predictive sensory edge': 3.4285714285714284,\n",
" 'distributed information systems': 2.5714285714285716,\n",
" 'distributed order management': 2.857142857142857,\n",
" 'e commerce': 3.2857142857142856,\n",
" 'e tailing': 3.2857142857142856,\n",
" 'm commerce': 3.2857142857142856,\n",
" 'edge computing': 3.2857142857142856,\n",
" 'ee paradigm': 2.5714285714285716,\n",
" 'encryption': 2.7142857142857144,\n",
" 'enterprise architecture': 2.571428571428571 6,\n",
" 'enterprise cloud': 4.142857142857143,\n",
" 'e learning': 3.5714285714285716,\n",
" 'electronic article surveillance': 3.142857142857143,\n",
" 'electronic governance': 2.5714285714285716,\n",
" 'eletronic product code': 2.5714285714285716,\n",

" 'epcis': 2.57,\n",
" 'eletronic product code information services': 2.5714285714285716,\n",
" 'embedded systems': 2.857142857142857,\n",
" 'emerging technologies': 2.857142857142857,\n",
" 'enterprise collaboration': 2.5714285714285716,\n",
" 'enterprise resource planning': 2.4285714285714284,\n",
" 'enterprise wide resource planning': 2.4285714285714284,\n",
" 'erp systems': 2.4285714285714284,\n",
" 'enterprise transformation': 3.2857142857142856,\n",
" 'equity crowdfunding platforms': 3.142857142857143,\n",
" 'ethereum': 2.8333333333333335,\n",
" 'facial recognition': 4.142857142857143,\n",
" 'fourth industrial revolution': 3.2857142857142856,\n",
" 'gamification': 3.4285714285714284,\n",
" 'general digital maturity': 3.2857142857142856,\n",
" 'geo fencing': 3.0,\n",
" 'gig economy': 2.4285714285714284,\n",
" 'global data synchronization network': 2.5714285714285716,\n",
" 'hackathon': 3.2857142857142856,\n",
" 'hadoop': 3.857142857142857,\n",
" 'high tech': 2.0,\n",
" 'horizontal integration': 3.142857142857143,\n",
" 'vertical integration': 3.142857142857143,\n",
" 'human machine interface': 3.7142857142857144,\n",
" 'hmi': 3.7142857142857144,\n",
" 'hybrid cloud': 4.0,\n",
" 'hyper converged technology': 3.2857142857142856,\n",
" 'hyper relevance': 2.5714285714285716,\n",
" 'iaas': 3.7142857142857144,\n",
" 'infrastructure as a service': 3.7142857142857144,\n",
" 'identity and access management': 3.5714285714285716,\n",
" 'image analytics': 4.142857142857143,\n",
" 'video analytics': 4.142857142857143,\n",
" 'indoor positioning system': 2.7142857142857144,\n",
" 'industrial internet of things': 4.428571428571429,\n",
" 'iiot': 4.428571428571429,\n",
" 'industry 4 0': 3.4285714285714284,\n",
" 'industry technologies': 2.142857142857143,\n",
" 'information security': 3.2857142857142856,\n",
" 'innovation infrastructure': 2.5714285714285716,\n",
" 'innovation processes': 2.5714285714285716,\n",
" 'innovation diffusion': 2.5714285714285716,\n",
" 'innovative it projects': 3.0,\n",
" 'integrated supply chain': 3.0,\n",
" 'integrated workplace management system': 3.142857142857143,\n",
" 'intelligent business services and related': 3.0,\n",

" 'intelligent devices': 3.857142857142857,\n",
" 'interactive screen': 4.0,\n",
" 'interactive voice response': 4.0,\n",
" 'ivr': 4.0,\n",
" 'internet of things': 4.571428571428571,\n",
" 'iot': 4.571428571428571,\n",
" 'internet technologies': 3.2857142857142856,\n",
" 'it automation': 3.5714285714285716,\n",
" 'itg mechanisms': 2.5714285714285716,\n",
" 'journey analytics': 3.142857142857143,\n",
" 'journey map': 3.142857142857143,\n",
" 'keyword stuffing': 3.1666666666666665,\n",
" 'knowledge intensive': 2.2857142857142856,\n",
" 'knowledge networks': 2.5714285714285716,\n",
" 'knowledge sharing': 2.7142857142857144,\n",
" 'level of digitalization': 3.0,\n",
" 'li fi': 2.857142857142857,\n",
" 'light fidelity': 2.857142857142857,\n",
" 'lights out environment': 2.142857142857143,\n",
" 'lights out manufacturing': 2.142857142857143,\n",
" 'linked open data': 2.2857142857142856,\n",
" 'lte': 2.142857142857143,\n",
" 'long term evolution': 2.142857142857143,\n",
" 'machine learning': 4.285714285714286,\n",
" 'machine to machine': 3.2857142857142856,\n",
" 'marketing automation': 3.4285714285714284,\n",
" 'maturity model': 2.5714285714285716,\n",
" 'micro blog': 2.857142857142857,\n",
" 'micro services': 3.857142857142857,\n",
" 'middleware': 2.7142857142857144,\n",
" 'mixed reality': 3.3333333333333335,\n",
" 'mobile apps': 4.166666666666667,\n",
" 'mobile devices': 4.285714285714286,\n",
" 'mobile health': 3.2857142857142856,\n",
" 'mobile technologies': 4.142857142857143,\n",
" 'mobile pos': 4.142857142857143,\n",
" 'mobile shopping': 4.142857142857143,\n",
" 'mobile wallet': 4.142857142857143,\n",
" 'modern information technologies': 2.0,\n",
" 'multi factor authentication': 3.142857142857143,\n",
" 'natural language processing': 4.428571428571429,\n",
" 'near field communication': 2.5714285714285716,\n",
" 'nfc': 2.5714285714285716,\n",
" 'network analytics': 3.4285714285714284,\n",
" 'network analysis': 3.4285714285714284,\n",
" 'network virtualization': 3.5714285714285716,\n",

" 'omni channel retail': 4.714285714285714,\n",
" 'one click check out': 3.4285714285714284,\n",
" 'online marketplace': 4.0,\n",
" 'online order management system': 4.0,\n",
" 'online touchpoint': 4.0,\n",
" 'open innovation': 2.2857142857142856,\n",
" 'open source software': 3.1666666666666665,\n",
" 'open source': 3.142857142857143,\n",
" 'open data': 3.142857142857143,\n",
" 'open systems interconnection model': 3.142857142857143,\n",
" 'osim': 3.142857142857143,\n",
" 'paas': 4.0,\n",
" 'integrated platform as a service': 4.0,\n",
" 'personalized marketing': 4.142857142857143,\n",
" 'personlized shopping experience': 4.142857142857143,\n",
" 'platform based': 2.142857142857143,\n",
" 'plug   work': 1.7142857142857142,\n",
" 'predictive analysis': 4.285714285714286,\n",
" 'predictive analytics': 4.285714285714286,\n",
" 'predictive modeling': 4.285714285714286,\n",
" 'predictive maintenance': 3.857142857142857,\n",
" 'prescriptive analysis': 4.0,\n",
" 'prescriptive analytics': 4.0,\n",
" 'pricing optimization': 3.5714285714285716,\n",
" 'pricing algorithm': 3.5714285714285716,\n",
" 'process innovation': 3.4285714285714284,\n",
" 'product information management': 2.857142857142857,\n",
" 'qr code': 3.4285714285714284,\n",
" 'quantum computing': 3.5714285714285716,\n",
" 'ransomware': 3.7142857142857144,\n",
" 'real time': 4.0,\n",
" 'realtime data': 4.0,\n",
" 'reference models': 2.4285714285714284,\n",
" 'resilient run time environments': 3.4285714285714284,\n",
" 'rfid': 3.2857142857142856,\n",
" 'radio frequency identification': 3.2857142857142856,\n",
" 'robotic process automation': 4.285714285714286,\n",
" 'rpa': 4.285714285714286,\n",
" 'saas': 3.857142857142857,\n",
" 'software as a service': 3.857142857142857,\n",
" 'scaled agile framework': 3.4285714285714284,\n",
" 'scan and pay': 3.142857142857143,\n",
" 'search engine marketing': 3.5714285714285716,\n",
" 'search engine optimization': 3.5714285714285716,\n",
" 'self checkout': 4.0,\n",
" 'self serve': 4.0,\n",

" 'semantic and syntactic interoperability': 3.0,\n",
" 'server message block': 2.4285714285714284,\n",
" 'server virtualization': 3.5714285714285716,\n",
" 'service innovation': 3.2857142857142856,\n",
" 'service oriented enterprise architectures': 3.5714285714285716,\n",
" 'smart devices': 4.285714285714286,\n",
" 'smart factory': 4.0,\n",
" 'smart manufacturing': 4.0,\n",
" 'smart governance': 3.7142857142857144,\n",
" 'smart meter': 3.857142857142857,\n",
" 'smart packaging': 3.857142857142857,\n",
" 'smart shelf': 3.857142857142857,\n",
" 'smart poster': 3.857142857142857,\n",
" 'smart price tag': 3.857142857142857,\n",
" 'smart mobility': 4.0,\n",
" 'smart products': 3.857142857142857,\n",
" 'smart retail': 4.285714285714286,\n",
" 'smart store': 4.285714285714286,\n",
" 'smart city': 4.285714285714286,\n",
" 'smart village': 4.285714285714286,\n",
" 'smart service': 4.0,\n",
" 'smart technologies': 4.0,\n",
" 'social commerce': 3.5714285714285716,\n",
" 'social media': 4.285714285714286,\n",
" 'social media marketing': 4.285714285714286,\n",
" 'social network analysis': 4.0,\n",
" 'social networking': 3.7142857142857144,\n",
" 'social networks': 3.7142857142857144,\n",
" 'social wifi': 3.0,\n",
" 'software defined wide area network': 2.7142857142857144,\n",
" 'sd wan': 2.7142857142857144,\n",
" 'software development': 2.0,\n",
" 'software engineering': 2.0,\n",
" 'software systems': 2.0,\n",
" 'software implementation': 2.0,\n",
" 'storage area network': 2.0,\n",
" 'streaming analytics': 4.0,\n",
" 'supervised learning': 3.2857142857142856,\n",
" 'task automation': 3.7142857142857144,\n",
" 'technology development': 3.142857142857143,\n",
" 'technological development': 3.142857142857143,\n",
" 'technology adoption': 3.0,\n",
" 'technology architecture': 2.857142857142857,\n",
" 'technology driven': 3.4285714285714284,\n",
" 'tokenization': 3.4285714285714284,\n",
" 'uav': 4.0,\n",

```
    " 'unmanned aerial vehicle': 4.0,\n",
    " 'drone': 4.0,\n",
    " 'unified commerce platform': 3.4285714285714284,\n",
    " 'unified retail': 3.4285714285714284,\n",
    " 'user centered': 3.2857142857142856,\n",
    " 'user experience': 4.0,\n",
    " 'user interface design': 4.0,\n",
    " 'ux ui': 4.0,\n",
    " 'user onboarding': 4.0,\n",
    " 'utility analytics': 2.857142857142857,\n",
    " 'virtual and augmented reality': 4.428571428571429,\n",
    " 'virtual reality': 4.428571428571429,\n",
    " 'augmented reality': 4.428571428571429,\n",
    " 'vr ar': 4.428571428571429,\n",
    " 'virtual assistant': 4.0,\n",
    " 'virtual machine': 3.857142857142857,\n",
    " 'virtual network': 3.5714285714285716,\n",
    " 'virtual omnichannel inventory management': 3.857142857142857,\n",
    " 'virtual world': 3.857142857142857,\n",
    " 'virtualization of consulting services': 3.142857142857143,\n",
    " 'visible light communication': 3.7142857142857144,\n",
    " 'vlc': 3.7142857142857144,\n",
    " 'visual recognition system': 3.7142857142857144,\n",
    " 'voice commerce': 3.0,\n",
    " 'voice controlled ai': 3.857142857142857,\n",
    " 'conversational ai': 3.857142857142857,\n",
    " 'voice search': 4.0,\n",
    " 'voice activated assistant': 4.0,\n",
    " 'video content': 2.857142857142857,\n",
    " 'warehouse management system': 3.4285714285714284,\n",
    " 'wms': 3.4285714285714284,\n",
    " 'wearable technology': 3.4285714285714284,\n",
    " 'web applications': 3.0,\n",
    " 'webrooming': 2.5714285714285716,\n",
    " 'website navigation': 2.5714285714285716,\n",
    " 'work design': 1.7142857142857142}"
   ]
  },
  "execution_count": 8,
  "metadata": {},
  "output_type": "execute_result"
 }
],
"source": [
 "kw_dict = dict(zip(keywords_file['Keywords Modified'],keywords_file['Rating1']))\n",
 "kw_dict"
```

```
   ]
  },
  {
   "cell_type": "code",
   "execution_count": 9,
   "metadata": {},
   "outputs": [],
   "source": [
    "import os\n",
    "path = 'E:\\\\GTA work\\\\Tushar\\\\Report Reading\\\\Annual Reports Missing\\\\SnP500-
Missed'"
   ]
  },
  {
   "cell_type": "code",
   "execution_count": 10,
   "metadata": {},
   "outputs": [],
   "source": [
    "from tika import parser\n",
    "import re\n",
    "from datetime import datetime"
   ]
  },
  {
   "cell_type": "code",
   "execution_count": 11,
   "metadata": {
    "scrolled": true
   },
   "outputs": [
    {
     "name": "stdout",
     "output_type": "stream",
     "text": [
      "adsk   2011\n",
      "\n",
      "akam   2011\n",
      "\n",
      "amat   2011\n",
      "\n",
      "amzn   2011\n",
      "\n",
      "anss   2011\n",
      "\n",
      "cdns   2011\n",
```

```
"\n",
"chrw   2011\n",
"\n",
"csco   2011\n",
"\n",
"dltr   2011\n",
"\n",
"ebay   2011\n",
"\n",
"eqix   2011\n",
"\n",
"expd   2011\n",
"\n",
"expe   2011\n",
"\n",
"fast   2010\n",
"\n",
"fast   2011\n",
"\n",
"fitb   2011\n",
"\n",
"flir   2011\n",
"\n",
"fosl   2009\n",
"\n",
"fosl   2011\n",
"\n",
"grmn   2010\n",
"\n",
"grmn   2011\n",
"\n",
"idxx   2011\n",
"\n",
"mdlz   2011\n",
"\n",
"nflx   2008\n",
"\n",
"nflx   2009\n",
"\n",
"nflx   2010\n",
"\n",
"nflx   2011\n",
"\n",
"nktr   2011\n",
"\n",
"ntrs   2011\n",
```

"\n",
"orly 2011\n",
"\n",
"swks 2010\n",
"\n",
"ads 2011\n",
"\n",
"aig 2011\n",
"\n",
"all 2011\n",
"\n",
"ame 2011\n",
"\n",
"amt 2011\n",
"\n",
"aos 2011\n",
"\n",
"aph 2011\n",
"\n",
"axp 2011\n",
"\n",
"a 2011\n",
"\n",
"chk 2011\n",
"\n",
"cl 2011\n",
"\n",
"cma 2011\n",
"\n",
"cms 2011\n",
"\n",
"cnc 2011\n",
"\n",
"cnp 2011\n",
"\n",
"cnx 2011\n",
"\n",
"cof 2011\n",
"\n",
"cvs 2011\n",
"\n",
"cvx 2011\n",
"\n",
"cvx 2013\n",
"\n",
"dfs 2011\n",

```
 "\n",
 "dnr   2010\n",
 "\n",
 "dnr   2011\n",
 "\n",
 "dov   2011\n",
 "\n",
 "dre   2011\n",
 "\n",
 "eix   2011\n",
 "\n",
 "emn   2009\n",
 "\n",
 "emn   2011\n",
 "\n",
 "ew   2011\n",
 "\n",
 "exc   2011\n",
 "\n",
 "fcx   2011\n",
 "\n",
 "fhn   2011\n",
 "\n",
 "fmc   2011\n",
 "\n",
 "gd   2011\n",
 "\n",
 "ge   2009\n"
]
},
{
"name": "stderr",
"output_type": "stream",
"text": [
 "2020-07-04 19:09:47,284 [MainThread  ] [WARNI]  Tika server returned status: 422\n"
]
},
{
"name": "stdout",
"output_type": "stream",
"text": [
 "'content'\n",
 "ge   2011\n",
 "\n",
 "gme   2011\n",
 "\n",
```

```
 "gnw   2011\n",
 "\n",
 "hbi   2011\n",
 "\n",
 "hog   2011\n",
 "\n",
 "hpq   2011\n",
 "\n",
 "hum   2011\n",
 "\n",
 "ipg   2011\n",
 "\n",
 "itw   2011\n",
 "\n",
 "jcp   2011\n",
 "\n",
 "jec   2011\n",
 "\n",
 "jef   2011\n",
 "\n",
 "jnj   2011\n",
 "\n",
 "jpm   2011\n",
 "\n",
 "kmx   2011\n",
 "\n",
 "ksu   2011\n",
 "\n",
 "k   2011\n",
 "\n",
 "lh   2011\n",
 "\n",
 "lm   2011\n",
 "\n",
 "mkc   2011\n",
 "\n",
 "mon   2011\n",
 "\n",
 "mtb   2011\n",
 "\n",
 "mtd   2011\n"
 ]
},
{
 "name": "stderr",
 "output_type": "stream",
```

```
 "text": [
  "2020-07-04 19:11:52,268 [MainThread  ] [WARNI]  Tika server returned status: 422\n"
 ]
},
{
 "name": "stdout",
 "output_type": "stream",
 "text": [
  "\n",
  "m   2011\n",
  "'content'\n",
  "nlsn   2010\n",
  "\n",
  "phm   2011\n",
  "\n",
  "psa   2011\n",
  "\n",
  "pwr   2011\n",
  "\n",
  "qep   2008\n",
  "\n",
  "qep   2009\n",
  "\n",
  "re   2011\n",
  "\n",
  "rig   2011\n",
  "\n",
  "rop   2011\n",
  "\n",
  "shw   2011\n",
  "\n",
  "stz   2011\n",
  "\n",
  "vtr   2010\n",
  "\n",
  "vtr   2011\n",
  "\n",
  "wmb   2011\n",
  "\n",
  "wmt   2010\n",
  "\n",
  "xom   2011\n",
  "\n",
  "xom   2015\n",
  "\n",
  "xxii   2008\n",
```

```
    "\n",
    "xxii   2009\n",
    "\n",
    "xxii   2010\n",
    "\n",
    "xxii   2017\n",
    "\n"
   ]
  }
 ],
 "source": [
  "company_dict = {}\n",
  "file_missed = []\n",
  "doc_len = []\n",
  "start_time = datetime.now()\n",
  "for path,dirs,files in os.walk(path):\n",
  "        content_str = "\n",
  "        for file in files:\n",
  "            word_dict = {}\n",
  "            try:\n",
  "                file_info = file.split('_')\n",
  "                company_ticker = file_info[1].lower()\n",
  "                year = file_info[2].split('.')[0]\n",
  "                if int(year) > 2007 and int(year) < 2018:\n",
  "                    company_year = company_ticker + '_' + year\n",
  "                    print(company_ticker, ' ',year)\n",
  "                    file_path = os.path.join(path,file)\n",
  "                    parsedPDF = parser.from_file(file_path)\n",
  "                    content = parsedPDF['content']\n",
  "                    content_edit = content.translate(table)\n",
  "                    content_edit1 = content_edit.lower()\n",
  "                    content_list = [word for word in content_edit.strip().split() if word.isalpha()]\n",
  "                    for key,val in kw_dict.items():\n",
  "                        counter = sum(1 for _ in re.finditer(r'\\b%s\\b' % re.escape(key),
content_edit1))\n",
  "                        word_dict[key] = counter * val\n",
  "                    word_dict['Word Length'] = len(content_list)\n",
  "                    company_dict[company_year] = word_dict\n",
  "                    print()\n",
  "            except Exception as e:\n",
  "                print(e)\n",
  "                file_missed.append(file)    \n",
  "                continue\n",
  "                \n",
  "  "end_time = datetime.now()       "
 ]
```

```
},
{
 "cell_type": "code",
 "execution_count": 12,
 "metadata": {},
 "outputs": [
  {
   "name": "stdout",
   "output_type": "stream",
   "text": [
    "0:09:15.900976\n"
   ]
  }
 ],
 "source": [
  "time_diff = end_time - start_time\n",
  "print(time_diff)"
 ]
},
{
 "cell_type": "code",
 "execution_count": 13,
 "metadata": {},
 "outputs": [],
 "source": [
  "company_df = pd.DataFrame(company_dict).T"
 ]
},
{
 "cell_type": "code",
 "execution_count": 14,
 "metadata": {},
 "outputs": [
  {
   "data": {
    "text/html": [
     "<div>\n",
     "<style scoped>\n",
     "    .dataframe tbody tr th:only-of-type {\n",
     "        vertical-align: middle;\n",
     "    }\n",
     "\n",
     "    .dataframe tbody tr th {\n",
     "        vertical-align: top;\n",
     "    }\n",
     "\n",
```

```
"    .dataframe thead th {\n",
"        text-align: right;\n",
"    }\n",
"</style>\n",
"<table border=\"1\" class=\"dataframe\">\n",
"  <thead>\n",
"    <tr style=\"text-align: right;\">\n",
"      <th></th>\n",
"      <th>3d planogram</th>\n",
"      <th>planogram automation</th>\n",
"      <th>3d space management</th>\n",
"      <th>adaptive integrated ea framework</th>\n",
"      <th>additive manufacturing</th>\n",
"      <th>3d printing</th>\n",
"      <th>advanced analytics</th>\n",
"      <th>agile methods</th>\n",
"      <th>agile development</th>\n",
"      <th>amazon retail analytics</th>\n",
"      <th>...</th>\n",
"      <th>voice activated assistant</th>\n",
"      <th>video content</th>\n",
"      <th>warehouse management system</th>\n",
"      <th>wms</th>\n",
"      <th>wearable technology</th>\n",
"      <th>web applications</th>\n",
"      <th>webrooming</th>\n",
"      <th>website navigation</th>\n",
"      <th>work design</th>\n",
"      <th>Word Length</th>\n",
"    </tr>\n",
"  </thead>\n",
"  <tbody>\n",
"    <tr>\n",
"      <th>adsk_2011</th>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>...</td>\n",
"      <td>0.0</td>\n",
```

```
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>75874.0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>akam_2011</th>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>...</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>48715.0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>amat_2011</th>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
```

```
"    <td>...</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>56291.0</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>amzn_2011</th>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>...</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>36947.0</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>anss_2011</th>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
```

```
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>...</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>46041.0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>...</th>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"     <td>...</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>xom_2015</th>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
```

```
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>...</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>16439.0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>xxii_2008</th>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>...</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>11241.0</td>\n",
"   </tr>\n",
"   <tr>\n",
"     <th>xxii_2009</th>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
"     <td>0.0</td>\n",
```

```
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>...</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>10790.0</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>xxii_2010</th>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>...</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
"    <td>18504.0</td>\n",
"  </tr>\n",
"  <tr>\n",
"    <th>xxii_2017</th>\n",
"    <td>0.0</td>\n",
"    <td>0.0</td>\n",
```

```
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>...</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>0.0</td>\n",
"      <td>17534.0</td>\n",
"    </tr>\n",
"  </tbody>\n",
"</table>\n",
"<p>109 rows Ã— 437 columns</p>\n",
"</div>"
],
"text/plain": [
"           3d planogram  planogram automation  3d space management  \\\n",
"adsk_2011           0.0                   0.0                  0.0   \n",
"akam_2011           0.0                   0.0                  0.0   \n",
"amat_2011           0.0                   0.0                  0.0   \n",
"amzn_2011           0.0                   0.0                  0.0   \n",
"anss_2011           0.0                   0.0                  0.0   \n",
"...                 ...                   ...                  ...   \n",
"xom_2015            0.0                   0.0                  0.0   \n",
"xxii_2008           0.0                   0.0                  0.0   \n",
"xxii_2009           0.0                   0.0                  0.0   \n",
"xxii_2010           0.0                   0.0                  0.0   \n",
"xxii_2017           0.0                   0.0                  0.0   \n",
"\n",
"           adaptive integrated ea framework  additive manufacturing  \\\n",
"adsk_2011                               0.0                     0.0   \n",
"akam_2011                               0.0                     0.0   \n",
"amat_2011                               0.0                     0.0   \n",
"amzn_2011                               0.0                     0.0   \n",
"anss_2011                               0.0                     0.0   \n",
"...                                     ...                     ...   \n",
```

```
"xom_2015                    0.0                0.0   \n",
"xxii_2008                   0.0                0.0   \n",
"xxii_2009                   0.0                0.0   \n",
"xxii_2010                   0.0                0.0   \n",
"xxii_2017                   0.0                0.0   \n",
"\n",
"          3d printing  advanced analytics  agile methods  agile development  \\\n",
"adsk_2011         0.0                0.0          0.0               0.0   \n",
"akam_2011         0.0                0.0          0.0               0.0   \n",
"amat_2011         0.0                0.0          0.0               0.0   \n",
"amzn_2011         0.0                0.0          0.0               0.0   \n",
"anss_2011         0.0                0.0          0.0               0.0   \n",
"...               ...                ...          ...               ...   \n",
"xom_2015          0.0                0.0          0.0               0.0   \n",
"xxii_2008         0.0                0.0          0.0               0.0   \n",
"xxii_2009         0.0                0.0          0.0               0.0   \n",
"xxii_2010         0.0                0.0          0.0               0.0   \n",
"xxii_2017         0.0                0.0          0.0               0.0   \n",
"\n",
"          amazon retail analytics  ...  voice activated assistant  \\\n",
"adsk_2011                      0.0  ...                        0.0   \n",
"akam_2011                      0.0  ...                        0.0   \n",
"amat_2011                      0.0  ...                        0.0   \n",
"amzn_2011                      0.0  ...                        0.0   \n",
"anss_2011                      0.0  ...                        0.0   \n",
"...                            ...  ...                        ...   \n",
"xom_2015                       0.0  ...                        0.0   \n",
"xxii_2008                      0.0  ...                        0.0   \n",
"xxii_2009                      0.0  ...                        0.0   \n",
"xxii_2010                      0.0  ...                        0.0   \n",
"xxii_2017                      0.0  ...                        0.0   \n",
"\n",
"          video content  warehouse management system  wms  \\\n",
"adsk_2011          0.0                          0.0  0.0   \n",
"akam_2011          0.0                          0.0  0.0   \n",
"amat_2011          0.0                          0.0  0.0   \n",
"amzn_2011          0.0                          0.0  0.0   \n",
"anss_2011          0.0                          0.0  0.0   \n",
"...                ...                          ...  ...   \n",
"xom_2015           0.0                          0.0  0.0   \n",
"xxii_2008          0.0                          0.0  0.0   \n",
"xxii_2009          0.0                          0.0  0.0   \n",
"xxii_2010          0.0                          0.0  0.0   \n",
"xxii_2017          0.0                          0.0  0.0   \n",
"\n",
"          wearable technology  web applications  webrooming  \\\n",
```

```
      "adsk_2011                0.0             0.0        0.0   \n",
      "akam_2011                 0.0             0.0        0.0   \n",
      "amat_2011                 0.0             0.0        0.0   \n",
      "amzn_2011                 0.0             0.0        0.0   \n",
      "anss_2011                 0.0             0.0        0.0   \n",
      "...                       ...             ...        ...   \n",
      "xom_2015                  0.0             0.0        0.0   \n",
      "xxii_2008                 0.0             0.0        0.0   \n",
      "xxii_2009                 0.0             0.0        0.0   \n",
      "xxii_2010                 0.0             0.0        0.0   \n",
      "xxii_2017                 0.0             0.0        0.0   \n",
      "\n",
      "            website navigation  work design  Word Length  \n",
      "adsk_2011                 0.0         0.0      75874.0  \n",
      "akam_2011                  0.0         0.0      48715.0  \n",
      "amat_2011                 0.0         0.0      56291.0  \n",
      "amzn_2011                  0.0         0.0      36947.0  \n",
      "anss_2011                 0.0         0.0      46041.0  \n",
      "...                       ...         ...          ...  \n",
      "xom_2015                  0.0         0.0      16439.0  \n",
      "xxii_2008                 0.0         0.0      11241.0  \n",
      "xxii_2009                 0.0         0.0      10790.0  \n",
      "xxii_2010                 0.0         0.0      18504.0  \n",
      "xxii_2017                 0.0         0.0      17534.0  \n",
      "\n",
      "[109 rows x 437 columns]"
     ]
    },
    "execution_count": 14,
    "metadata": {},
    "output_type": "execute_result"
   }
  ],
  "source": [
   "company_df"
  ]
 },
 {
  "cell_type": "code",
  "execution_count": 15,
  "metadata": {},
  "outputs": [],
  "source": [
   "company_df.to_excel('E:\\\\GTA work\\\\Tushar\\\\Report
Reading\\\\AnualReportInfo_Rating_2008_to_2018_missed.xlsx')"
  ]
```

```
},
{
 "cell_type": "code",
 "execution_count": 16,
 "metadata": {},
 "outputs": [
  {
   "data": {
    "text/plain": [
     "['NYSE_GE_2009.pdf', 'NYSE_M_2011.pdf']"
    ]
   },
   "execution_count": 16,
   "metadata": {},
   "output_type": "execute_result"
  }
 ],
 "source": [
  "file_missed"
 ]
},
{
 "cell_type": "code",
 "execution_count": 17,
 "metadata": {},
 "outputs": [
  {
   "data": {
    "text/plain": [
     "2"
    ]
   },
   "execution_count": 17,
   "metadata": {},
   "output_type": "execute_result"
  }
 ],
 "source": [
  "len(file_missed)"
 ]
},
{
 "cell_type": "code",
 "execution_count": null,
 "metadata": {},
 "outputs": [],
```

```
  "source": []
 }
],
"metadata": {
 "kernelspec": {
  "display_name": "Python 3",
  "language": "python",
  "name": "python3"
 },
 "language_info": {
  "codemirror_mode": {
   "name": "ipython",
   "version": 3
  },
  "file_extension": ".py",
  "mimetype": "text/x-python",
  "name": "python",
  "nbconvert_exporter": "python",
  "pygments_lexer": "ipython3",
  "version": "3.7.0"
 }
},
"nbformat": 4,
"nbformat_minor": 2
}
```

TABLE 1: Extract of the output obtained from keywords analysis of annual reports of companies

Company Ticker	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	Grand Total
MON	0.00	0.00	0.00	0.00	0.00	19.43	19.29	28.29	52.14		119.14
A	33.57	36.29	35.14	54.57	51.60	55.43	27.86	36.29	86.74	90.88	508.36
AAP	19.43	52.29	61.29	83.00	89.86	86.00	67.43	54.43	54.43	23.71	591.86
AAPL	152.00	155.43	202.24	197.83	145.29	187.57	196.43	132.86	193.00	233.71	1796.36
ABC	7.43	49.14	6.14	12.57	49.43	12.29	11.43	0.00	3.71	7.00	159.14
ABMD	0.00	4.86	2.00	6.29	5.86	9.86	8.57	4.71	12.43	12.14	66.71
ABT	25.14	3.43	3.71	15.17	26.48	25.45	28.88	54.88	38.31	37.17	258.62
ACN	176.29	183.14	36.57	47.14	36.86	51.00	16.43	49.71	227.31	75.00	899.45
ADBE	180.86	315.14	611.86	1029.98	1281.55	859.38	916.02	924.31	965.95	736.00	7821.05
ADP	34.43	33.57	41.43	69.86	89.14	75.29	97.14	84.14	79.57	74.71	679.29
ADS	109.29	91.00	136.14	211.29	292.43	286.57	282.57	211.86	191.29	165.71	1978.14
ADSK	69.00	160.29	104.14	104.14	174.57	144.57	152.14	177.86	290.14	225.00	1601.86
AEE	10.14	6.14	11.43	14.00	4.00	14.86	15.43	22.29	41.86	29.57	169.71
AES	0.00	0.00	17.71	36.29	28.43	30.29	25.43	37.43	32.29	32.71	240.57
AET	0.00	4.00	12.00		33.43	32.29	43.71	97.86	128.86		352.14
AFL	10.86	6.14	2.86		11.57	7.71	0.00	0.00	7.29	22.14	68.57
AGN	0.00	0.00	0.00	19.29	7.14	70.86	71.00	70.00	46.14	62.71	347.14
AIG	20.29	20.29	20.00	20.00	7.57	11.29	23.00	68.86	32.71	115.29	339.29
AIV	9.71	25.29	21.29	10.00	0.00	4.29		0.00	7.86	16.43	94.86
AIZ	0.00	0.00	0.00	12.00	19.43	32.29	28.00	36.57	20.00	77.29	225.57
AJG	0.00	0.00	7.14	7.14	7.14	14.43	10.57	9.14	16.43	28.86	100.86
AKAM	77.57	178.43	198.29	222.60	165.14	159.29	231.00	237.60	0.00	7.57	1477.48
AKS	0.00	0.00	3.43		6.29	0.00	7.14	0.00	0.00	3.43	20.29
ALGN	29.14	32.29	59.57	26.00	31.71	50.57	61.14	22.86	30.29	22.29	365.86
ALL	0.00	7.29	9.14	2.00	6.86	47.29	47.29	59.57	77.71	92.14	349.29
ALXN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	19.43	49.29	0.00	68.71
AMAT	0.00	61.71	0.00	62.86	95.57	59.86	76.71	76.57	99.29	125.71	658.29
AMD	68.00	68.00	56.14	103.43	98.00	76.71	130.86	88.74	94.43	155.40	939.71

TABLE 2

Descriptive Information About Industries And Companies Within Industries

NAICS3Dgt	Industry Sub-sector Description	No Of Cos In Sample DataSet	No Of Total Indy Cos Available in Master
111	Crop Production	1	12
211	Oil & Gas Extraction	20	226
212	Mining and quarrying (except oil and gas)	3	147
213	Support activities for mining, and oil and gas extraction	3	66
221	Utilities	18	153
236	Construction of Buildings	3	34
237	Heavy and Civil Engineering Construction	1	42
238	Specialty trade contractors for contruction activities	1	15
311	Food Manufacturing	6	93
312	Beverage and Tobacco Product Manufacturing	6	45
314	Textile Product Mills	1	5
315	Apparel Manufacturing	2	42
321	Wood Product Manufacturing	1	22
322	Paper Manufacturing	3	41
323	Printing and related support activities	1	18
324	Petroleum and Coal Products Manufacturing	4	50
325	Chemical Manufacturing	30	993
326	Plastics and Rubber Products Manufacturing	1	33
331	Primary Metal Manufacturing	1	58
332	Fabricated Metal Product Manufacturing	2	70
333	Machinery Manufacturing	14	193
334	Computer and Electronic Product Manufacturing	46	713
335	Electrical equipment, appliance and component manufacturing	3	97
336	Transportation equipment manufacturing	13	135
337	Furniture and related product manufacturing	1	23
339	Miscellaneous manufacturing (toys, jewellery, sporting goods, office supplies)	12	179
423	Merchant Wholesalers, Durable Goods	4	101
424	Merchant Wholesalers, Nondurable Goods	4	78
425	Wholesale Electronic Markets and Agents and Brokers	1	2
441	Motor vehicle and parts dealers	3	22
442	Furniture and home furnishings stores	1	10

443	Electronics and appliance stores	2	11
444	Building material and garden equipment and supplies dealers	1	6
446	Health and personal care stores	3	25
448	Clothing and Clothing Accessories Stores	3	60
452	General Merchandise Stores	7	24
453	Miscellaneous Store Retailers	1	8
454	Nonstore Retailers	1	53
481	Air Transportation	1	33
482	Rail Transportation	2	9
483	Water transportation	2	75
484	Truck Transportation	2	26
486	Pipeline Transportation	2	69
488	Support activities for transportation	1	18
492	Couriers and messengers	2	7
511	Publishing Industries (except Internet)	10	278
512	Motion Picture and Sound Recording Industries	1	26
515	Broadcasting (except Internet)	8	63
517	Telecommunications	1	149
518	Data Processing, Hosting, and Related Services	8	89
519	Other Information Services	7	253
522	Credit Intermediation and Related Activities	21	822
523	Securities, Commodity Contracts, and Other Financial Investments and Related Activities	16	209
524	Insurance Carriers and Related Activities	23	190
531	Real estate	22	287
532	Rental and Leasing Services	1	43
541	Professional, Scientific, and Technical Services	11	270
561	Administrative and Support Services	4	96
562	Waste Management and Remediation Services	2	23
611	Educational Services	2	43
621	Ambulatory Health Care Services	4	88
721	Accommodation services	3	36
722	Food services and drinking places	3	82
812	Personal and Laundry Services	1	15
999	Non-classifiable Establishments	1	156
	TOTAL	389	7360

TABLE 3: Descriptive Statistics (Raw values – non-standardized)

Variables	Obs	Mean	Std.Dev.	Min	Max	p1	p99	Skew.	Kurt.
Firm_Age (In years)	3887	69.496	49.944	3	233	6	211	0.859	2.841
Firm_Sales~t (In $Million)	3731	17518.2	38306	3.919	483521	297.06	175752	6.603	63.15
EnvmDynm_S~s (Environmental Dynamism)	3731	0.03	0.017	0.009	0.19	0.01	0.084	2.068	10.02
ScoretoWor~o (Digital Transformation Score)	3643	0.001	0.002	0	0.034	0	0.012	5.236	43.33
TMT_Role_H~y (Role Heterogeneity)	3233	55.824	43.754	1	386	9	234	2.521	12.32
TMT_MeanTe~e (Mean tenure in days)	3224	1338.38	536.721	17	4211.667	258.73	3025	0.985	4.918
CEO_Total_~e (Total tenure in days)	3241	3710.78	2319.69	-418	13879	200	11027	1.143	4.698
CEO_Pwr1_D~u (CEO duality)	3272	0.573	0.495	0	1	0	1	-0.29	1.086
CEO_Pwr2_S~u (Shareholding percentage)	2945	1.227	3.821	0	53.834	0.005	20.731	7.098	69.29
CEO_Pwr3_F~u (Founder Status)	3272	0.026	0.158	0	1	0	1	5.998	36.98

TABLE 4 - Correlations
Pairwise correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Digital Transformation	1.000										
(2) Firm_Age	-0.137*	1.000									
(3) Firm_Sales	0.110*	0.245*	1.000								
(4) Fiscal Year	0.265*	-0.001	0.100*	1.000							
(5) Environmental Dynamism	0.049*	-0.100*	0.130*	-0.133*	1.000						
(6) Role Heterogeneity	0.206*	0.267*	0.572*	0.259*	-0.051*	1.000					
(7) TMT Mean Tenure	-0.058*	0.002	-0.050*	0.068*	-0.094*	-0.079*	1.000				
(8) CEO Total Tenure	-0.079*	-0.105*	-0.152*	-0.229*	0.055*	-0.110*	0.192*	1.000			
(9) CEO Power (Duality)	-0.135*	0.222*	0.172*	-0.029	-0.019	0.112*	0.071*	0.259*	1.000		
(10) CEO Power (Shareholding)	-0.042*	-0.241*	-0.375*	-0.209*	0.120*	-0.309*	0.076*	0.483*	0.117*	1.000	
(11) CEO Power (Founder Status)	0.063*	-0.128*	-0.089*	-0.040*	0.043*	0.058*	0.061*	0.193*	-0.043*	0.265*	1.000

* shows significance at the .05 level

TABLE 4A : Model building for 3-level model

VARIABLES	Comparison: Null Model (1)	Comparison: Null Model with Level 2 RI (Firm) (2)	Comparison: Null Model with Level 2 and Level 3 RIs for Firm and Industry (3)	Comparison: Model 1 - FI, TS, Level 2 and Level 3 RIs (4)	Comparison: Model 2 - FI, TS, Level 2 RI and RS, Level 3 RI (5)	Estimation: Model 3 - CVs, TS, Level 2 RI and RS, Level 3 RI (6)	Estimation: Model 4 - All Variables, TS, Level 2 RI & RS, Level 3 RI (7)
	Model_000	Model_00	Model_0	Model_1	Model_2	Model_3	Model_4
Fixed Effects:							
Fiscal Year				0.169***	0.173***	0.165***	0.153***
				`(0.006)	(0.007)	(0.008)	(0.011)
Firm_Age						-0.003**	-0.004**
						(0.001)	(0.001)
Firm_Sales						0.153***	0.061
						(0.042)	(0.055)
Role Heterogeneity							0.230**
							(0.101)
TMT Mean Tenure							0.701
							(5.025)
TMT Mean Tenure* TMT Mean Tenure							-0.035
							(0.330)
CEO Total Tenure							-0.073
							(0.111)
CEO Power (Duality)							-0.057
							(0.070)
CEO Power (Shareholding)							0.022
							(0.024)
CEO Power (Founder Status)							-0.387
							(0.248)
Constant	3.528***	3.524***	3.390***	2.462***	2.444***	1.326***	-1.251
	(0.031)	(0.074)	(0.157)	(0.161)	(0.162)	(0.394)	(19.093)

VARIABLES	Comparison: Null Model (1)	Comparison: Null Model with Level 2 RI (Firm) (2)	Comparison: Null Model with Level 2 and Level 3 RIs for Firm and Industry (3)	Comparison: Model 1 - FI, TS, Level 2 and Level 3 RIs (4)	Comparison: Model 2 - FI, TS, Level 2 RI and RS, Level 3 RI (5)	Estimation: Model 3 - CVs, TS, Level 2 RI and RS, Level 3 RI (6)	Estimation: Model 4 - All Variables, TS, Level 2 RI & RS, Level 3 RI (7)
	Model_000	Model_00	Model_0	Model_1	Model_2	Model_3	Model_4
Random Effects:							
Level 3 - var(_Cons)			1.136***	1.134***	1.140***	1.065***	1.076***
			`(0.284)	`(0.284)	`(0.286)	`(0.269)	`(0.296)
Level 2 - var(ED)					5421.818**	5314.403**	4793.087**
					`(1167.737)	`(1152.541)	`(1651.125)
Level 2 - var(_Cons)		1.900***	0.945**	0.968**	0.993**	0.970**	0.961**
		`(0.151)	`(0.089)	`(0.089)	`(0.090)	`(0.088)	`(0.096)
Level 2 - var(Residual)	3.390***	1.500***	1.496***	1.237***	1.083***	1.082***	1.039***
	-0.079	`(0.037)	`(0.037)	`(0.030)	`(0.030)	`(0.030)	`(0.034)
Model Fit:							
LR-Likelihood	`-7395.277	`-6385.058	`-6317.672	`-6010	`-5805.17	`-5804.816	`-4403.648
Wald Chi2 (DF)	Null	Null	Null	686.12(1)***	551.99(1)***	570.67(3)***	418.74(10)***
Observations	3,643	3,643	3,643	3,643	3,554	3,554	2,729
Number of groups		370	65	65	65	65	55

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

RI = Random Intercept

FI = Fixed Intercept

TS = Time Slope

RS = Random Slope

CV = Control Variables

AV = All Variables

MCHD = Mean Centered High Dynamism

MCLD = Mean Centered Low Dynamism

TABLE 4B : Model building for 2-level model

VARIABLES	Comparison: Null Model (Same as Level 3 Model)	Comparison: Null Model with L2 RI (Industry)	Comparison: Model 1 with FI, TS and L2 RI (Industry)	Estimation: Model 2 with CVs, TS, Level 2 RI & RS (Indy)	Estimation: Model 3 with All Variables, TS, Level 2 RI & RS (Indy)
	`(1)	(2)	(3)	(4)	(5)
	Model_000	Model_00	Model_1	Model_2	Model_3
Fixed Effects:					
Year_Id			0.169***	0.166***	0.164***
			(0.008)	(0.010)	(0.011)
Firm_Age				-0.003***	-0.003***
				(0.001)	(0.001)
ln_FirmSales_CRSP_Cmpstat				0.153***	0.114***
				(0.023)	(0.033)
Ln_Role_Hetero					0.067
					(0.061)
Ln_TMT_MeanTenure					9.353*
					(4.779)
Ln_TMT_MeanTenure* Ln_TMT_MeanTenure					-0.633**
					(0.313)
ln_0_CEO_TotalTenure_0					-0.135
					(0.086)
CEO_Pwr1_Duality_ExeCompu					-0.104*
					(0.063)
Log_CEO_Pwr2_ShrOwnPrcntx100					0.050**
					(0.022)
CEO_Pwr3_FndrStat_ExeCompu					-0.092
					(0.183)
Constant	3.528***	3.352***	2.421***	1.231***	-32.022*
	(0.031)	(0.163)	(0.169)	(0.257)	(18.205)

VARIABLES	Comparison: Null Model (Same as Level 3 Model)	Comparison: Null Model with L2 RI (Industry)	Comparison: Model 1 with FI, TS and L2 RI (Industry)	Estimation: Model 2 with CVs, TS, Level 2 RI & RS (Indy)	Estimation: Model 3 with All Variables, TS, Level 2 RI & RS (Indy)
	`(1)	(2)	(3)	(4)	(5)
	Model_000	Model_00	Model_1	Model_2	Model_3
Random Effects:					
Level 2 - var(ED)				488.874	135.495
				`(315.071)	`(157.383)
Level 2 - var(_Cons)		1.610**	1.626***	1.535**	1.581**
		`(0.309)	`(0.309)	`(0.293)	`(0.326)
Level 2 - var(Residual)	3.390***	2.285***	2.046***	1.996***	1.914***
	`(0.079)	`(0.054)	`(0.048)	`(0.048)	`(0.053)
Model Fit:					
LR-Likelihood	`-7395.277	-6775.337	`-6581.213	`-6398.649	`-4848.686
Wald Chi2 (DF)	Null	Null	418.63(1)***	387.47(3)***	335.51(10)***
Observations	3,643	3643	3,643	3,554	2,729
Number of groups		65	65	65	55

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

RI = Random Intercept
FI = Fixed Intercept
TS = Time Slope
RS = Random Slope
CV = Control Variables
AV = All Variables
MCHD = Mean Centered High Dynamism
MCLD = Mean Centered Low Dynamism

TABLE 5A: 3-level full model direct effects and moderation testing results

Hypothesis	3 level hierarchy	Direct	Moderation - Hi_CenED	Moderation - Lo_CenED	Moderation - Hi_Binary	Moderation - Lo_Binary
H1a, b	TMT role heterogeneity	0.230(0.02)	0.256(0.06)	0.144(NS)	0.307(0.02)	0.127(NS)
H2a,b	TMT mean tenure	0.700(NS)	-2.11(NS)	11.639(0.09)	-5.13(NS)	4.89(NS)
H2a,b	TMT mean tenure squared	-0.035(NS)	0.142(NS)	-0.761(0.09)	0.336(NS)	-0.301(NS)
H3a	CEO Power - Duality	-0.057(NS)	-0.083(NS)	-0.050(NS)	-0.106(NS)	-0.043(NS)
H3b	CEO Power - Shareholding	0.022(NS)	0.021(NS)	-0.005(NS)	-0.003(NS)	0.035(NS)
H3c	CEO Power - Founder	-0.387(NS)	-0.073(NS)	-0.579(0.06)	-0.668(0.03)	-0.141(NS)
H4a,b	CEO Total Tenure	-0.073(NS)	-0.733(NS)	-0.094(NS)	0.064(NS)	-0.147(NS)

TABLE 5B: 2-level full model direct effects and moderation testing results

Hypothesis	2 level hierarchy	Direct	Moderation - Hi_CenED	Moderation - Lo_CenED	Moderation - Hi_Binary	Moderation - Lo_Binary
H1a, b	TMT role heterogeneity	0.067(NS)	0.173(NS)	0.012(NS)	0.225(0.00)	-0.088(NS)
H2a,b	TMT mean tenure	9.353(0.05)	-0.283(NS)	20.04(0.00)	1.115(NS)	16.787(0.01)
H2a,b	TMT mean tenure squared	-0.633(0.04)	-0.003(NS)	-1.331(0.00)	-0.104(NS)	-1.11(0.01)
H3a	CEO Power - Duality	-0.104(0.09)	-0.198(0.07)	-0.061(NS)	-0.097(NS)	-0.152(0.09)
H3b	CEO Power - Shareholding	0.049(0.02)	0.040(NS)	0.059(0.03)	0.035(NS)	0.058(0.08)
H3c	CEO Power - Founder	-0.092(NS)	0.082(NS)	-0.281(NS)	0.183(NS)	-0.308(NS)
H4a,b	CEO Total Tenure	-0.135(NS)	-0.039(NS)	-0.176(0.093)	-0.009(NS)	-0.207(0.09)

For both, table A and B, bold values indicate significance at 0.05 level, underlined italics indicate significance at 0.1 level, and NS means non-significant. ED = Environmental Dynamism; Hi CenED = High Dynamism on mean centered ED variable; Lo CenED = Low Dynamism on mean-centered ED variable