2016

Selling the Fourth Revolution: The Correlation between C-Suite Architecture and a Big Data Mindset as Portrayed in the Letter to the Shareholders

Samuel Weinberg
Wharton, UPenn

Selling the Fourth Revolution: The Correlation between C-Suite Architecture and a Big Data Mindset as Portrayed in the Letter to the Shareholders

Abstract

Big data's diverse applications for the modern data deluge span problems and industries. While offering titular possibilities, *is big data an area of serious corporate inquiry or is it a source of hype?*

This research seeks to add to the growing body of management literature on big data. C-suite architecture additions of a CIO and/or a CTO demonstrate an environment for and increased BDA mindset. A review of the annual letter to the shareholder is a proxy for the external narrative of a big data strategy.

To measure correlation between C-suite structure and outward narrative, a regression for eight industries, from 2011 to 2014, measure correlation between external and organizational positioning. Additional regressions examine industry-sensitivity and leader-laggard dynamics. Five of the eight industries demonstrate correlation between architectural repositioning and perceived investor support for big data, indicating that big data intensive C-suite architecture correlates with letter to the shareholder big data emphasis.

Keywords

Big Data, C-Suite, Letter to the Shareholder, Organizational Positioning, External Positioning

Disciplines

Business

Selling the Fourth Revolution: The Correlation between C-Suite Architecture and a Big Data Mindset as Portrayed in the Letter to the Shareholders

Samuel Weinberg
College of Arts and Sciences
wsam@sas.upenn.edu

Thesis for The Wharton School at the University of Pennsylvania
April 27, 2016

Advised by Dr. Rahul Kapoor, Associate Professor of Management, Wharton School of the University of Pennsylvania; kapoorr@wharton.edu

Acknowledgements
To Whitney Kramer for answering my many ambiguous questions, Dr. Utsav Schurmans for encouraging me to pursue the Wharton Research Scholars Program, Professor Catherine Schrand, for distilling what was then a very nebulous idea and pointing me in the direction I needed to go, and Dr. Rahul Kapoor for his unending patience, guidance when my process began to stray, and continual demonstration of what scholarly research should be, I could not have completed this without you all.

Table of Contents

Terms and Abbreviations

Big Data: the umbrella of emerging technologies that include analytics, the Cloud, the Internet of Things, data algorithms, and information-based cybersecurity

Big Data Analytics: the analytic and algorithmic sorting and scrapping of larges data series

Broken Data / Dark Data / Inactionable: data that is, in and of itself, meaningless (e.g. misspelled Twitter messages)

B2B: Business-to-Business interactions

B2C: Business-to-Customer interactions

CAO: Chief Analytics Officer

CDO: Chief Data Officer

CIO: Chief Information Officer (for management research, CIO is never the Chief Investment Officer)

CMO: Chief Marketing Officer

CSO: Chief Security Officer

C-Suite: C-suite level upper management, beneath the Chief Executive Officer

CTO: Chief Technology Officer

External Position: the narrative a firm expresses publically or to shareholders about its relationship with a potential technology or strategy

HADOOP: open-source software framework for storing data and running applications on clusters of commodity software[1]

Organizational Position: C-Suite architecture relating to officers in context to other positions

IoT: Internet of Things, the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment[2]

LSH: Letter to Shareholders, the document presented to shareholders on an annual basis by the CEO

NAICS: North American Industry Classification System

[1] SAS Hadoop definition
[2] Gartner IoT definition

Abstract

Big data's diverse applications for the modern data deluge span problems and industries. While offering titular possibilities, *is big data an area of serious corporate inquiry or is it a source of hype?*

This research seeks to add to the growing body of management literature on big data. C-suite architecture additions of a CIO and/or a CTO demonstrate an environment for and increased BDA mindset. A review of the annual letter to the shareholder is a proxy for the external narrative of a big data strategy.

To measure correlation between C-suite structure and outward narrative, a regression for eight industries, from 2011 to 2014, measure correlation between external and organizational positioning. Additional regressions examine industry-sensitivity and leader-laggard dynamics. Five of the eight industries demonstrate correlation between architectural repositioning and perceived investor support for big data, indicating that big data intensive C-suite architecture correlates with letter to the shareholder big data emphasis.

Key words

Big Data

C-suite

Letter to the Shareholder

Organizational Positioning

External Positioning

Introduction

Analyzing big data is a diverse and flexible information solution that spans from understanding broken information to machine communication. Beginning as an analysis of information technology (IT) expansion, big data initially was not considered separate from standard data, but rather an expansion (Agneeswaran 2012). Early research did not distinguish big data from that which already existed. However, as capabilities improved, research classified big data in its own niche. Now, as big data enters disparate industries, research at a management level has gained traction.

This research seeks to add to the body of management research on big data. Specifically, this analysis explores whether or not C-suite restructuring correlates with big data presentation to shareholders as demonstrated in the letter to the shareholder (LSH). For consistency, this research uses a dataset of firms that are shored in the United States or Canada. Thus, omissions are made, especially in pharmaceutical manufacturing where many industry leaders are shored in the United Kingdom.

Additional data parameters also define the dataset. This research examines firms and industries for the past four years of recorded letter to the shareholder data, from 2011 to 2014. The following analysis assesses four hypotheses via multivariate regressions. The regressions demonstrate correlation between an organization’s C-suite structure and its external positioning of big data adoption as demonstrated through the LSH. Ultimately, the research indicates that a positive correlation between C-suite structure and external positioning exists on a select basis.

Hypotheses

H1a: *Firms with higher positioned C-suite big data officers (i.e., technology and IT related officers in the C-suite) display a relatively more intense external demonstration of a big data mindset in the letter to the shareholder.*

H1b: Firms with higher positioned C-suite big data officers (i.e., technology and IT related officers in the C-suite) do not display a relatively more intense external demonstration of a big data mindset in the letter to the shareholder.

H2: Firms with higher positioned C-suite big data officers (i.e., technology and IT related officers in the C-suite) display a relatively more intense external demonstration of a big data mindset in the letter to the shareholder on an industry-specific basis.

H3: Firms with higher positioned C-suite big data officers (i.e., technology and IT related officers in the C-suite) display a relatively more intense external demonstration of a big data mindset in the letter to the shareholder by income on an industry-specific basis.

Big data began as a major trend at the turn of the decade. In the past five years, there has been enough research to identify relevant industries for analysis. Within these industries, large firms often offer a greater field of analysis and represent larger industry trends (Davenport and Dyche 2013). Further, since accessible research trends exist and big data promises are both nuanced and apparent, additional variables are incorporated in this research. Net income is assumed as a proxy for firm ranking within its industry, and due to various claims of big data as an optimizer, number of employees is weighed in context to net income. The relationship between income and employees, while not the primary correlation, demonstrates other claims made by big data's offerings.

This research question surveys the historical examination of big data as a unique trend, and it examines thought on the digital revolution as a basis of analysis for further innovation. First, an examination of warehousing evolution explores the capacity for maintaining more data, and thus, its role as a catalyst for further trends. Dovetailing open-source capabilities is cybersecurity, since its needs multiply across otherwise unsecured networks. Further, the history explores more front-of-mind trends, such as analytics and algorithmic big data techniques. Big data is lastly explored in context to the Internet of Things (IoT). Since the multifarious big data applications exist in different industries, this research analyzes trends themselves as well as the specific firm-level application. Additionally, this research contextualizes the C-suite structure via an exploration of its history and the changing value of the data-related C-suite positions. Finally, a general history investigates letters to the shareholder as a firm's demonstration of external position and changes over time.

The literature review first assesses the initial gestating dialogue on big data. An understanding of the early narrative gives insight into initial thoughts on the trend. Next, the research analyzes literature on big data today and its function as the backbone of strategy-selection on the industry level. Literature on the various trend technologies has recently been published for each industry included in the research. Given the relative nature of this research, it also explores literature on the C-suite and LSH as a predictive instrument. In context to the former, the CIO is given especial attention. Finally, the literature review also examines research into past big data studies focused on the CIO and CTO.

The methodology section explains the logic behind each research decision. First, the analysis explores the decision to examine big data through the context of the C-suite. It does this by outlining the expanding nature of relatively new roles.[3] Next, the industry selections are discussed on the bases of diversity and germaneness. These are relative to both the facets of big data and the nature of an industry to be either business-to-business (B2B) or business-

[3]These include the CIO, CTO, CDO, CAO, CSO and on a firm-dependent basis the CMO

to-consumer (B2C). Finally, an in-depth explanation and analysis of the rationale of the LSH is examined.

Big data’s ascent is a trend that is too large to assume that no radiating changes within organizations will occur, and this study gives some indication that this has already occurred within C-suite architecture of some industries. When refined to five particular industries, the second hypothesis is validated while the other three are proven null. The sentiment is one of highly placed big data officers correlated with a sense of shareholder enthusiasm.

Historical Background

In the 1980s, firms' adoption of digital products transformed from infeasible to necessary. Numerous productive limitations (Sterling 1997) as well as professions disappeared overnight in the wake of this new trend (Chaudhuri 2009). Firms went from gradual incorporation of digital technologies to a unified fever-pitched adoption, and the Internet and digitization propelled bids over which firms would become leaders and laggards (Betsy 1999). As dramatic as the digital revolution was, it represented the beginning of a more pivotal big data revolution.

Much as the Second Industrial Revolution grew out of the First, are the promises of the Fourth meant to solve the limitations of the Third? The Big Data Revolution, at its core, alleges to make data more actionable, communicative, and less expensive (Agneeswaran 2012). Many firms did not begin discussing the potentials of big data until after the turn of the decade. Most industries, historically treating IT as a necessity, pivoted and now recognize the opportunity to benefit from heightened information capabilities (Barnes 2012). While the information itself may not be monetized, a paradigm shift in value is occurring.

For these firms with a developing mindset, big data can unlock new information opportunities, such as affordable storage or nuanced security. Whereas warehousing large datasets has historically been cost-prohibitive, Cloud computing now allows firms to decrease warehousing costs (Assuncao 2015). In the past, a company could have spent millions of dollars on hardware and physical space to warehouse far less information than they now can at a fraction of the cost. A terabyte of conventionally stored data costs $37,000 to maintain but on the Cloud the same terabyte is reduced to a $5,000 maintenance (Barnes 2012). Cloud computing, with the promise of higher-capacity and cheaper data storage, is actually the precursor to subsequent opportunities that newly obtainable storage can provide.

In weighing big data's major changes to how industries act, all previous limitations must be accounted for. The impact that the Cloud has on cheaper storage, however, is not enough to up-end industries around the world. Another reason firms have historically limited their data storage is noise (De Simoni, Judah and Zaidi 2015). Previously, most collectable data was perceived as inactionable and labeled as either noise, broken, or dark data. However, new technologies and algorithmic programs such as HADOOP make previously useless data actionable and potentially even profitable (Huang and Laney 2015). On a surface level, more data can simply be recorded. While conversations and clothing tried on while in-store were previously lost, sophisticated technologies combined with the Cloud can capture virtually all interactions. Recordings, footage, and streams of text can now be aggregated and processed through sophisticated algorithms (Huang and Laney). Currently, Facebook, Google, and their cohort companies have begun to mine these databases for forecasting trends (Burns, Jin and Friedman 2014). However, big data analytics forecasting capabilities may be spreading throughout very different industries and throughout the world.

Finally, in the current stage of data storage and machine learning algorithms, the IoT has become a realized possibility. The IoT, the third major promise of big data, is the capacity for machines to communicate over a veritable second Internet via micro-sensors, machine-learning, and extensive algorithms (Bradley, Ng and Thibodeau 2014). All three of these sub-trends have major consequences. An optimization could lead to positive returns via optimization that also could eliminate an entire class of employees. A pertinent question, beyond this speculation, is which firms, if any, plan to utilize big data to its full capacity?

A firm can express a big data mindset internally as well as externally. The internal demonstration of position, the C-suite, has seen radical changes since the 1990s as firms grapple with emerging technologies and needs (Guadalupe, Li and Wulf 2014). Particular to big data is the CIO, in charge of information systems and IT. As IT becomes more important and if more firms adopt a big data mindset, the CIO and IT will necessarily become more central in the firm's architecture (Aileen and Cummings 2016). IT was previously seen as a basic maintenance service in many industries, but now, it has the potential to become the veritable firm-wide laboratory for big data.

In addition to the CIO, other C-suite level positions have risen in prominence. Some firms dedicate a CMO or CAO to marketing or analytics. Further the CTO has become a central role as has, for many firms, the CDO. A somewhat common trend is for the CDO or CTO to act instead of a CIO in the big data role; it is rare for a firm to have both positions as redundancy could occur (Gerth and Peppard 2016). [4] When firms do adopt an aggressive architecture, it can be with a number of different strategic profiles. At a basic level, this is usually in the form of a high-ranking CIO or CTO role and potentially another high order position of similar consequence.

As the external evidence of a firm's strategic profile, the LSH can demonstrate the stance a firm is taking on its big data mindset to investors.[5] While usually under 1500 words, the LSH is a means for a firm to pursue a specific image to maintain confidence from its investors. The letter to the shareholder demonstrates a firm's desired reputation (Geppert and Lawrence 2008). If the letter to the shareholder is big data intensive, it is reasonable to postulate that the firm is putting a strong foot forward in its big data claim.

A caveat of a firm's LSH is the degree to which the document is constrained by its respective industry. Figure One demonstrates that commercial banking letters are sprawling documents while home healthcare letters are often much more concise. A pharmaceutical manufacturing document is often rich in laboratory conclusions, and an advertising piece will often be unavailable due to the privatized nature of the industry. Therefore, the industry

[4] Lexus Nexus Data.
[5] McKinsey developed the term 'big data mindset'.

narrative, as derived from the LSH, is a valuable, if imperfect, item of measurement for external effort.

Literature Review

As with most major trends, big data began on a small scale and grew. Initially, research on big data began as either speculative strategy or in reference to computer science. Big data first became a conceptual possibility in the late 2000s when scholars began to assess the direction that big data acceleration could take industries (Chaudhuri 2009). Many of these initial proposed strategies were directly related to information management and therefore applied most to the data-science industry. Further analyses investigated the roadblocks this offering initially ran into and produced needed software developments (Venkatachalam 2014). During these early years, big data and management theory remained separate.

A turning point for big data as an addition to the strategy mix occurred when data-intensive firms began marketing potential solutions provided by this offering. SAP notably has offered management training in elevating IT for several years, and large management consulting firms have also published literature on achieving a big data mindset (Kohers 2015). A big data solutions arms race has occurred in management consulting literature as many firms sought to become the trend leader. Accenture took a position of assessing success from early big data strategy positions (Accenture 2014). Deloitte showcased similar early trends and areas of research but did so in a more skeptical manner (Deloitte 2015). McKinsey, which adopted a position of cautious authority and recommended its brand of adoption on a case-by-case basis, did a third form of review a year earlier (McKinsey 2013). Universally, these firms have recommended strategies to begin elevating IT capabilities as well as applying open-sourcing opportunities from the Cloud without compromising security.

A sizeable body of literature has recently been written on a myriad of big data applications for each of the industries assessed in this research. At a general level, cybersecurity literature tied to big data has been explored, with especial attention to problems raised by the open-sourcing necessary for the Cloud (Patel 2015). General analytics research has also been heavily explored (Duncan, Linden, Koehler-Kruener and Zaidi 2015). While analytics is not new, a previously unobtainable level of sophistication is required to understand the deluge of data now available. A subdivision of analytics literature is specifically related to the opportunities dictated by open-sourcing (Duncan and Laney 2015). These examinations form the contextual backbone for understanding the industry-specific applications of the Cloud.

Some industry literature is more narrowly focused on explaining technologies in context to existing systems. Two such industry examples of big data related analytics are in banking (Morris, Sun, Xie, Xu and Zhu 2014) and insurance (Wade 2012). Retail (Burt, Davison, Hetu and Marian 2014) and healthcare (Cribbs and Handler 2015) literature on analytics and algorithms are weighted toward the exploration of aggregating previously broken data into actionable information. Literature also explores potential new avenues for firms pursuing information strategy mixes previously outside of their industry. One example is mobile applications for previously technology-sparse industries such as general grocers (Bartel and Scheibenreif 2014). Not only are these applications allowing firms to incorporate more sophisticated strategies, but literature highlights that these ‘apps’ also become vehicles themselves for further data collection.

Another blossoming frontier for big data has been the veritable second Internet that exists for machine-to-machine communication. IoT literature has examined the many diverse uses for micro-sensor networks and the vast algorithms that organize them and allow for communication. One application in IoT for stationary pipelines is smart-grid communication, which applies to the pipeline industry examined later in this research (Chen et al 2014). This microsystem can detect leaks or other potential security threats at the most minute level. While specific to the industries selected, the potentials and applications for the IoT eclipse the body of research studying them.

With any large technological trend, the human side of industry transformation is explored in tandem with actual advances. One controversial claim of human literature on big data is the displacement of a portion of the current workforce. Scholarly studies, such as the MIT Review, elucidate where employment may give way to pure automation (Rotman 2013). At first, blue-collar jobs such as manufacturing would further erode. As trends progress, white-collar occupations such as airline pilots could also be replaced by algorithmic programs and complex IoT systems (2013). Recently, the Federal Trade Commission released a report warning of potential dangers of transformations in the job market wrought by these trends (FTC Report 2016). On an individual basis, firms will have to decide how best to implement strategies in big data and weigh in the looming body of literature on the potentially destabilizing effects of this technology trend.

The recent surge in research on big data, both high-level and specific, has been accompanied by management-related strategy assertions. With CIO and IT elevation, recent management theory in this space has explored the potential dangers the CIO faces when attempting this massive overhaul (Gerth and Peppard 2016). These sharp changes in the power structure theoretically could to reverberate throughout the C-suite. Further, extensive research has been done on how upper management grapples with major shifts and turning points in a firm's strategy (Cruz, Ginel and Ordaz 2015). On a one-off basis, the C-suite may be repositioned or expanded to meet the previously unexpected needs of a firm developing a big data mindset.

As the letter to the shareholder is not a recent development, there is an extensive body of literature on the subject. Scholars have found the LSH has questionable predictive validity (Fisher and Hu 1989). However, this does not discount the intrinsic value in studying the letter from a management research paradigm. The LSH demonstrates how a firm wants to frame its narrative to the shareholders (Geppert and Lawrence 2008). While past literature does not demonstrate that the LSH is a proxy for firm activity, it is an important piece of social commentary provided by the firm.

Little research exists on specific framing of big data to investors. There is not a large body of literature that examines the two-way correlation between shifts in C-suite architecture and positioning in the LSH relative to big data. As a result, research with either variable acting as the driver on the other, in context to big data, is limited. This research seeks to add new material to the ongoing contemplation of how far big data is permeating into US and Canadian firms. A correlation between organizational structure and external pursuit is one metric for how genuinely firms view the possibilities of big data. The industry-segmenting and leader-laggard sensitivities of the analysis adds definition to the correlation.

Methodology

The methodology for this research explores the productive analysis of the strategies and policies by which firm managements pursue structuring and externally communicate these maneuvers. Also, the methodology examines the best metrics to make quantitative analyses for these shifts. An exploration of the environment for the analysis identifies timeline germaneness as well as industries for analysis.

The newness of the big data trend demonstrates many dynamic opportunities as well as hurdles. Finding a common metric to assess the big data mindset, both within and between industries, posed a potential problem. A number of strategies were available, albeit with indirect information, for understanding the environment within and between industries.

A major shift is occurring in upper management to address the changing big data landscape, one that is not uniform between industries or even within. Thus, a discrepancy is forming in the means by which firms are developing their big data mindset. This presents an important question: does this specific C-suite restructuring occur in order to promote trends in data analytics or as a result of the trends? Also, do C-suites that are more adapted for a big data mindset position their claim to investors around it? By assessing a wide range of industries in context to these questions, correlation can either be established or proven null.

Many firms have relied on either the existing CIO or CTO to establish new information systems and strategies (Aileen and Cummings 2016). In addition, many companies have also adopted a CMO, CDO or CAO for strategy elevation. These officers, often coming from more information or data-driven backgrounds (CIO Review 2015), are given authority over data analytics projects and drive the gestating big data mindset forward (Jarvenpaa 1990).

Given the chimeric nature of this trend, it is necessary to establish discreet boundaries in which to analyze big data. First, bifurcation between industries that lend themselves to

B2B or B2C interplay should occur (CIO Review 2015). One major field of big data is dedicated to aggregating actionable data about consumer experiences and consumer co-purchases (Huang and Laney 2015). Another segmentation is utilization of the IoT. Certain industries, such as manufacturing, utilize massive machines that can benefit from the IoT (Brynjolfsson and McElheran 2015) whereas, for example, direct life insurance carriers benefit less from maintaining an intricate network of communication between autonomous devices (Harris-Ferrante 2016). Fortunately, a number of factors, such as having a large dataset or algorithms to effectively sort data, hold constant for all industries employing a big data mindset.

Supposing certain standards that make an industry qualified for this examination into correlation between C-suite shifts and big data mindset adoption, tangential factors must be accounted for. In selecting industries, the most pertinent factors are relevance and diversity. Selected industries should demonstrate a desire to adopt the big data mindset. Also, diversity in the different presupposed sets of B2B, B2C, and IoT relevance are necessary. As a result of these stipulations, eight industries were selected for the regression. With the spectrum of big data strategies accounted for, eight industries should provide adequate insight into structural changes at the managerial level.

In deciding industries for examination, it is critical to avoid an ad hoc approach. Specific firms across numerous industries would independently be interesting cases for analysis; however, a prevailing, or even moderate, big data mindset does not exist in the industry itself. As a result, this research used stringent standards to select eight industries. Similarly, industry selection based on percent of US GDP or 10K line items relating to revenue or research and development had immediate appeal due to its inherently systematic nature. To achieve a balance between a potpourri of interesting selections and one that is unnecessarily off target just to achieve rigor, the final list is comprised of reports from major consulting firms, existing research on the field, financial news, and information systems databases in the last five years.

When assessing viable industries, an important final delineation is anomaly verses norm. For example, within mining, promising examples exist but the mining industry as a whole has not displayed a trend toward a big data mindset (Latimore 2015). In developing a dialogue on big data analytics as a multi-industry trend, specific omissions were made, and the industries selected were intentionally not enshrined in data themselves; companies such as SAP or Google would not have illuminated how big data is expanding to the farthest spaces in the global economy (ACM 2014). Thus, given consulting reports, databases, scholarly reviewed journals, and financial news, the eight industries selected are presented with each respective NAICS code attached.

1. Commercial Banking (522110)
2. Pharmaceutical Preparation Manufacturing (325412)
3. Pipeline Transportation of Crude Oil (486110)
4. Home Healthcare Services (621610)
5. Men and Women's Clothing (4481(1/2)0)
6. Supermarkets and other Grocery Stores (445110)
7. Direct Life Insurance Carriers (524113)
8. Advertising Agencies (541810)

Having selected industries, the next step assesses C-Suite shifts from 2011-2014. The range of years selected establishes a period during which BDA first emerged as a prominent feature of the firm strategy. The dynamic of either rapid managerial shift or ease due to previous alignment could have been expressed. This analysis had the potential to express itself in tandem with the LSH shifts that may have occurred in the same timeframe. Industries also entered big data at varying rates, with multi-industry prominence occurring after the turn of the decade (Davenport and Dyche 2013). As a result, using any year prior to 2010 exposes the research to greater industry-specific third variables than are necessary.

Within the C-suite, a number of different interactions may have occurred during the given timespan for a select firm. Prevalence of CIO and CTO in the top 50 executives as well as movement within the hierarchy over the years implied value shifts. If CTO fluctuation had been extremely common within a given industry or between industries since 2010, volatility within the position and demand for elite data professionals would have been a potential extrapolation.

In addition to the chronological boundaries, this research explores a consistent number of companies per industry. Within each of the eight industries, individual firms were selected from the leading five to ten companies. Thematically, the number and attendance to only the most successful firms served two purposes. First, this limiting factor dealt with time constraints and viability of the research. Second, while the first and, potentially, second company in an industry could have displayed a best-of-breed tactic and mobilized higher disposable income, the fourth and fifth were more indicative of the overall industry trend.

Equally important is what a firm externally does with its strategy of engaging BDA. A firm could be highly invested in a big data mindset, but given past trends or industry trends, not present its BDA. Presentation in this instance pertains to investment-related presentation. Thus, the LSH can act as a proxy for externally motivated big data enthusiasm.

Studying the LSH allows for the quantitative analysis necessary for regressions. Across each year's LSH, the research analyzes the percent of the letter dedicated to big data. Various words to inform each word search were selected on an industry-specific basis as well as on a more general basis. The research compares percents on an industry level and total research basis, and a multivariate regression incorporating other variables parsed between variables with a correlation and those with a lack thereof.

In assessing the external big data mindset, certain analyses were made in the LSH. Once industries were identified by the previous denominations, language on the dataset size acted as a base for examination. As datasets have both grown and become cheaper to maintain (Sharma 2016), companies across industries have begun to compile sets that are as large as viably possible (Venkatachalam 2014). The largest driver of this exponential growth is the practice of containing broken data. A desire to hold this data is the first step for many companies toward developing their big data profile. Broken data does not need to be in the form of a B2C interaction. Many firms are attempting to capture minute time logs and use history for employee computers and specific programs. Therefore, this field of word analysis acts as the general word search applicable across all industries.

Alongside the general word search, a more nuanced search occurred for each industry. In home healthcare, certain customer management systems and application-based innovations may have occurred that are important to note but do not apply to other industries (Cribbs and Handler 2015). One potential danger of this was a weighting toward a higher percent of the LSH dedicated to big data in industries with more 'unique' words. However, the results of the quantitative analysis show that this reservation was not realized.

Upon collecting this information, the research analyzed C-suite architectural shifts in context to LSH content. Given the numerous other variables, a plethora of sensitivities were available. Where correlation exists, after sensitivity, a congruency was demonstrated in the big data mindset on an external and organizational structuring basis.

Case Studies

Case 1: Capital One – Utilizing Big Data to Break from the Pack

Commercial banking is at a major junction in the United States and Canada. As credit transformed commercial banking for employees and customers alike, big data has the potential to do the same nearly a century later. With data storage growth, large and small banks now require more intensive warehousing and text-based analytics programs (2016), and these initial needs have had a cascading effect on the industry. Hand-in-hand, banks are moving to the Cloud for more cost-effective storage, which has generated a need for more sophisticated cybersecurity and IT at CIOs' direction (Moyer 2016). In this wave of big data, not all bank strategies are equally positioned. Some firms are falling behind while others are readying themselves to be the leaders in a new age.

Capital One has built a legacy of bold innovation, leaping into new technologies within the credit card space and now big data. Internally, the firm has built a narrative for its success around the initial treatment of credit card technology as analogous to scientific innovation (Fairbank 1997), and it built a veritable laboratory to test new technologies for credit cards (2007). By experimenting with actual customer data, the bank began to develop like a technology firm, with trial-and-error research for each product. With this approach, Capital One enticed investors and customers alike, developing an impressive offering and eclipsing far older institutions (Andrews, Cheatham, Immaneni and McCombs 2007). Now, the tech-savvy bank is portraying itself as rewriting the banking process with big data. Thus, when compared to its closest competitors, Capital One makes an excellent case for analyzing the relationship between big data and the C-suite.

If there is indeed a correlation between BDA and LSH, Capital One's actions suggest it would have an industry-dominating mention of BDA in the LSH in tandem with a well-positioned CIO. Since 1998, Robert Alexander has been Capital One's CIO. His tenure has extended through the entirety of this research period, and he holds the company's third-rank position, behind the CEO and CFO. Further, from Figure Two, Capital One's BDA narrative does dominate the industry in context to its LSH, and in this space, the company has steadily moved ahead of its competition over the past four years.

While the C-suite makeup is IT-intensive and the LSH is an external discussion of this, is Capital One reaping the rewards of what it claims? Rather than comparing revenue and employment relative to the entire industry, Capital One's promises can best be compared against banks with similar revenue. In assessing Capital One's BDA profile, it is best to compare the firm to US Bancorp and PNC, the firms respectively ahead of and behind Capital One.

In comparing these three companies, the amount of each LSH dedicated to BDA shows a clear differentiation, and Capital One has demonstrated dedication to information by being the industry leader in CIO positioning. Therefore, Capital One can be examined via an inverse correlation between revenue and employment that is stronger than its competitors. This claim is grounded in the fact that predictive analytics and complex algorithms should, at least, not require a large employment expansion for relative revenue (Morris, Sun, Xie, Xu and Zhu 2014). As time has gone on, these trends should have steadily increased. Income should rise with stagnating or decreasing employment as new roles are asserted

Figure three demonstrates somewhat surprising results. While Capital One does follow patterns predicted by big data theory, it is has moved away from those predictions over time. A number of exogenous factors to big data could very well be causing this transition or this shift could be a blip in an otherwise highly predicted trend. Beyond acting as a test for a predictive model, Capital One represents a well-hedged big data strategy. As a bank that has the median revenue of the ten commercial banks included, big data is an opportunity for firms with a big data mindset to distance itself from its competition. This strategy has already led to higher shareholder interest and potential revenue growth relative to competitors (Andrews, Cheatham, Immaneni and McCombs 2007). It remains to be seen how Capital One's closest competitors will meet this aggressive positioning.

Capital One's big data story is one that applies outside of the world of commercial banking. Techniques in big data that allow for a commercial bank to be successful are also important for most other financial institutions (News RX 2011). These firms rely on nuanced relationships, understandings of the market, algorithms, and at a higher level, the ability to rapidly parse through large sets of information. The same techniques Capital One used could be useful to firms within the Direct Health Insurance industry (Harris-Ferrante 2016). As a result, Capital One's model proposes a strategy question for not just one industry, but instead, an entire sector.

Case 2: Crude Oil Transportation – Unboxing a Bewildering Industry Trend

The crude oil pipeline transportation industry is one of deftly navigating a complex landscape. Investors expect hefty dividends that are often above ten percent (Filbeck and Visscher 2003), while extensive regulations require official 'open-seasons' for expansion to occur (Hirsh and Spatding 2012). Given the intense regulation and spatially massive nature of pipelines, certain technological procedures are necessary. Pipelines have constantly developed new security and GPU measures to increase safety (Lambert 2015). Further, these technologies allow for better-optimized crude oil flow.

Despite the implicit difficulties in the crude oil pipeline transportation industry, companies still profit. This industry has seen high returns with very low employment, which maintains high competition despite numerous exogenous shocks (Latimore 2015). Thus, big data is a potential galvanizing point for a market leader like Enterprise Products Partners to beat out competitors. Big data could also be a tool of midsized companies like Magellan to adopt a Capital One-approach to differentiation.

There are a number of forecasted needs for big data within the crude oil transportation industry. With labyrinthine networks of pipeline, algorithms and IoT technologies could allow for GPU automated services (Eldred and McAvey 2014). Relative to other IoT potentials, this would allow for higher margins and lower risk and fewer ground workers. In line with legal constraints, constant security measures lead to fewer disastrous events and improved legal relations (Jensen and Perrons 2015). This is critical for positive public relations as well as the opportunity for more access to open seasons. Therefore, it is worth examining how firms are grappling with this potential.

Two immediate trends stand out within Figure Four for this industry. First, the CIO is either well placed or not present within the C-suite. Second, there is virtually no mention of big data in the LSH across the industry. For market leaders, such as Enterprise Products Partners and Kinder Morgan, this is a seemingly odd strategy. The third most important member of the C-suite for these industry leaders is not even tangentially present in the LSH. In examining the LSH, the documents are frequently over 1000 words, implying that concision is not correlative with this trend. The actual contents of these documents almost exclusively relate to revenue streams and open seasons, and they do very little to portray IT leaps within their firm or compared to their competitors.

In examining Figure Five, firms with a CIO/CTO in rank two and no mention of big data in the LSH, a few potential explanations appear. First, the financial sophistication and focus of investors is a possibility. Rather than differentiating firms by their big data mindset, they may be more affected by the very high dividends (Filbeck and Visscher 2003). These investors could even be deterred by the presence of seemingly needless information. A second explanation could be a tepid culture within the shareholder industry toward big data

(Jensen and Perrons 2015). It is possible that either or both of these causes contributed to the trend.

The trends within the pipeline industry act as an example for potentially larger issues implicit in examining the shareholder narrative as a proxy for external demonstration of a big data mindset. Certain firms or industries are not yet at the point where externally expressing their big data mindset is advantageous, and there is a litany of reasons for why this might be. The LSH is a metric to assess external big data dialogue. While a correlation between big C-suite architecture and LSH content may exist, there are other important variables in play.

Quantitative Analysis

Multivariate regressions assess potential correlations within the selected body of research. The overarching question of correlation is presented through the architectural shift via potential impacts it has on the narrative of the LSH. The LSH is made into two variables to contextualize the variability for percent; the division is along the length of the documents themselves and the percent of the documents that are dedicated to discussing big data-related concepts.[6] C-suite architecture is split into two variables as well. In order to delve deeper into potential architecture, second-order positions are included if at all present. This allows for correlation to be expressed via first-order C-suite architectural shifts and second-order if they exist.

Firm selection and time-spread give context for any potential trends. Both industry and year act as environmental variables of analysis. Rank-order percent LSH of firms is examined (see Figure 6). An additional regression is dedicated to a potential leader-laggard dynamic (see Figures 11). Net income acts as a proxy for rank as this could otherwise become a controversial claim.

In order to contextualize research within the existing literature, firm size is expressed within the context of big data promises. Big data should lead to large increases in IT (Lin and Weng 2014) and a general decrease in blue and even white-collar jobs (Rotman 2013). Therefore, the research includes the number of employees at each firm. If the number of employees decreases while net income increases in a given year, the proposed trend is being demonstrated. This additional variable influenced case selection. Certainly, there are many third variables that could be the cause as well for the inverse relationship between increasing income and decreasing employment.

---

[6] Captured in specific word-search, such that the total percent is actually higher when entire big data related sentences are counted in 'percent LSH'

The data in Figure Seven skews toward commercial banking with 38 examples. Clothing retail, pharmaceutical manufacturing and direct life insurance follow with 26, 27, and 27 examples, respectively. This is due to the nature of the industries as well as that many of the largest companies in these industries are shored in the United States or Canada. Other industries such as grocers, which have only 13 examples, are primarily shored in Europe or had one firm that is equity acquisitive of other players. Another cause of low example yield is privatization, which occurs most in Advertising, where there are only four examples.

Shifting strategies in presentation also affects frequency. All firms in commercial banking exhibit data for at least three years, whereas home healthcare firms, while often not private, do not have available information for every year. Firm prominence is a major cause of the discrepancy, as private firms could not be captured. The dichotomy between industries that present LSH and those that do not may speak to other trends in the research.

The dataset requires an addendum for interpretation. Due to constraints, both in terms of capturing data and time, and the set contains 165 items; it is possible that a statistical analysis with a larger dataset would find different conclusions. Also, causation cannot be extrapolated even with a very strong correlation result. Preponderance within certain industries also adds a potential skew to this research. Future research should expand upon these questions to strengthen scholarship in this area.

Given that the analysis examines C-suite architecture and LSH, certain industries are more easily captured than others. Figure One demonstrates C-suite architecture, net income and number of employees were not difficult to capture,[7] but different industries treat the LSH differently. This is both in context to content and availability. While varying content is useful for examining correlation, lack of availability for certain industries led to low representation. As a result, while purposefully diversified, the research gravitates in the direction of financial

[7] Necessary information is available via LexusNexus Corporate Affiliates

firm examination, with commercial banking and direct life insurance as the two largest industry representatives.

In Figure One, the average position across all eight industries for the big data officer (CIO/CTO or other) is quite low, averaging the 23.353th position in the C-suite. advertising, while somewhat negligible due to small sample, is highest with 6.25th, and crude oil transportation is lowest with 37.48th. The high for most industries is second, but for home healthcare, it is third and for advertising it is fourth.

The LSH yields similar results to the C-suite architecture. Figure One shows, across industries, on average, .0009% of the letter to the shareholder discusses big data-related concepts. Thus, a marginally larger percent is dedicated to these concepts as the analysis only counts exact words and not their entire sentences. Further, crude oil transportation and direct life insurance bring down the average significantly as they each have .00006% and .00003% of words respectively dedicated to big data.

The primary cause of the low position is the number of firms without a big data officer. When firms without an officer in the C-suite related to big data are removed in Figure Eight, there is not a change in sensitivity. Thus, the lack of officers does not impact the lack of correlation in Figure Nine. A potential explanation for this is the drastic decrease in the sample size as many firms do not have an officer regardless of industry.

Correlation is not proven via Figure Nine, the multivariate regression. T-statistics were quite high and a confidence interval is not established in the initial regression. At this point in time, neither a positive nor negative correlation exists between the percent of the LSH dedicated to big data and the C-suite architecture for the sample. Correlated shifts are also not suggested. There are a number of potential explanations for this result, and in order to examine them, a number of sensitivity tests need to be explored and followed by a new regression.

A potential sensitivity test removes industries with high propensities of no mention of big data in the LSH. This removal is based on the dependent rather than independent variable. The primary drawback of this approach is that a greater constraint is placed on the total sample size. While uncorrelated data could be removed, this would decrease the likelihood of a correlation expressing itself. Further, these industries do have firms with mention of big data. All industries within the study mention either the percent LSH or C-suite members who are big data related.

Upon removal of the three industries with the lowest percent LSH, a strong correlation manifests itself in Figure Ten. With C-suite architecture as the independent variable and percent LSH dedicated to big data as the dependent variable, a P-value of .030 occurs. Thus, a correlative relationship between C-suite architecture and the percent LSH occurs for commercial banking, pharmaceutical manufacturing, home healthcare, advertising, and clothing retail. A firm with a high-ranking officer who is either in charge of technology or IT is also correlatively more likely to demonstrate a big data mindset to the shareholders relative to the industry.

While industry correlation is established, industry position relative to correlation is a differentiating question. Industries are ranked in each year by net income and, as such, could be bifurcated between leaders and laggards. Here, it is worth noting if the CIO/CTO has a greater effect on the contents of the LSH if said firm is a leader within its industry. This regression incorporates the same variables as past ones but introducing the new binary leader-laggard dynamic.

The results of the quantitative analysis in Figure 11 demonstrate that leaders and laggards are no more likely than the other to have a C-suite architecture that is correlated with a dependent percent LSH dedicated to big data. The regression shown in Figure 8 examines the correlative effect between percent LSH dedicated to big data and firms without either a CIO or CTO. Therefore, correlations dictate that the environment for CIO figures to impact the external narrative to shareholders continues to be an open one.

Conclusion

Despite being a relatively recent development, the potential for big data has not gone unnoticed. While initially seen as a series of disparate technologies, big data has been labeled a singular trend, but that is not to say the potentials have been treated as such (De Prato and Simon 2015). Numerous applications, either general or industry-specific, have been touted as productivity optimizers and have been carefully researched in recent technical and business journals (McKinsey 2013). As a result, management theory has begun to examine big data within the larger body of theory and research.

Generally, the C-suite position most linked to big data and information technology is the CIO. The CIO is not a new position; management research has studied it for decades (Grover, Jeong, Kettinger and Lee 1993). However, the early roles for the CIO within most industries were less central to the dynamic growth of the firm. Since big data most often applies to the analysis of information, the CIO is usually the role most closely associated with onboarding these new technologies.[8] Thus, this research sought to analyze the role of the C-suite architecture on some aspect of a firm's big data image with attention most often attached to the CIO.

The word-capture procedure for the LSH offered a quantitative analysis of external positioning. Choosing this mode of analysis did not predict the future of a firm's data acquisitiveness (Fisher and Hu 1988). However, rather than assessing what a firm does, this research analyzed what a firm says. Thus, a big data-intensive letter to the shareholder demonstrates, first, that a firm is confident in its plan for big data implementation. Second, the firm feels its investors will not be put off by this technology investment and may even feel more positive toward the firm's direction as a result.

---

[8] Depending on industry, the CTO, CAO, CSO, CDO and CAO may be in charge of big data related specific or even general projects as well.

Despite stringent methodology, additional variables exist within the research. Industries were selected on a basis of relevancy and should, within the spread of the eight selected, capture the different potentials of big data. Further, these industries should engage in big data, but not include firms dedicated to data. Also, a diversity of B2B and B2C industries were used in the study, and a spread of four years, 2011 to 2014, limited the time period of the research. The research would have benefited from 2015 LSH information and C-suite information. However, due to time limitations, this was not viable. Net income existed as a tangible ranking attribute for firms within industries.

Additionally, the final data set used in this analysis did have further limitations. Due to LSH privacy, certain industries were scant. Many firms assessed within the Grocer industry were owned subsidiaries of large international grocery chains. Since a requirement of the data for homogenous C-suite analysis was to only include firms that are shored in the United States or Canada, the equity-acquisitive firms could not be used. Within the Advertising industry, similar issues pervaded. The majority of companies not owned by one of the big four advertising conglomerates are private.[9] This was such that almost all major advertising firms elected either private status or subsidiary status. As a result of these factors, a relatively small and financial-heavy dataset of 165 points came together.

This research deployed multiple regression-analyses. Sensitivities across industries and a firm's financial position within its industry were incorporated to expunge potential third variables. An initial regression demonstrated no correlation; with all eight industries incorporated, there was no significant correlation between C-suite architecture and percent LSH dedicated to big data. A second regression excluded firms without a big data-related officer as they could not demonstrate a correlation. These omissions of these firms further lowered the sample size, but no significant correlation was observed. A third regression

[9] WPP, Interpublic, Omnicom and Publicis Groupe

excluded crude oil transportation and direct life insurance, the two industries with the lowest percent of LSH dedicated to big data, and this variation did not result in any correlation. Finally, the next lowest industry for percent of LSH dedicated to big data, grocer, was removed as well.

Upon removing the three lowest industries, the findings indicate a correlation of 97% between organizational and external positioning. The findings demonstrate that the correlation between C-suite structuring and percent LSH dedicated to big data also had a T-statistic of 2.20. This represents a strong correlation – a 95 percent confidence interval – between organizational positioning (C-suite architecture) and external positioning (percent of LSH dedicated to big data). While the LSH may not be a superb document for predictive analysis or even definitive current analysis, it expresses correlation with C-suite structure for big data positioning.

With five industries demonstrating this trend, the research explored the discrepancy between industry leader and laggard in Figure 11. This regression demonstrated no correlation. At a correlative level, a firm's percent LSH dedicated to big data relative to C-suite architecture (and vice-versa) occurred regardless of position.[10] Also, this correlatively demonstrates that leaders are not outpacing laggards in either metric relative to the other.

A number of extrapolations from the correlations can be found. The results of this analysis extend to how different industries express big data motivation. First, firms are organizationally and externally interested in big data. When reversing perspective, shareholders are also interested in the potential for this trend. Shareholders view big data as a possibility for furthering a position, developing a new strategy, or remaining competitive.

Further research should continue the dialogue in this research area. By expanding the number of industries, number of firms within industries, or revisiting this same dataset with

[10] Within commercial banking, pharmaceutical manufacturing, men and women's clothing retail, advertising and home healthcare

additions from future years, a clearer correlation – or lack thereof – could become apparent. Further, a different metric could be correlated with C-suite architecture, and thus, other representations of external positioning could be explored.

One major consequence of big data is heavy automation. While not central to the research question, this poses potential problems for employees. Increased automation can lead to commodification of customers and desired expendability of employees (Behrend, Karim and Willford 2015). Since a correlation does not exist for leaders verse laggards in developing a big data mindset, firms are no less likely to have correlated organizational architecture and external shifts if they are large or small.

Across industries, firms are modernizing their C-suites and their external positions to shareholders accordingly as they relate to big data, and these decisions underpin a strategic move. It is a misnomer to relegate the label of a ‘big data mindset’ to Google and its tech-centered peers, or even to select industries. The world itself is developing a big data mindset, and the firm is in a transition.

Appendix

Figure 1: LSH /C-Suite 1 by industry with mean, median, mode, min

industry	numbe~ds	oflhs	csuite~1
Advertising Agen	528.75	.00885	42.75
	446	.0045	39
	636	.0177	46
	88.23217	.0061803	3.304038
Clothing Retail	1544.115	.0016154	16.59091
	568	0	0
	3089	.0047	48
	710.5493	.0015104	22.50219
Commercial Banki	4623.816	.0012289	31.97368
	668	0	0
	21573	.0125	48
	4637.313	.0022301	20.29177
Crude Oil Transp	1549.421	.0000632	12.53333
	535	0	0
	3216	.0012	48
	798.0429	.0002753	21.51699
Direct Life Insu	1542.593	.0000259	28.51852
	299	0	0
	3018	.0007	48
	739.7459	.0001347	22.37507
Grocer	1608.692	.0010231	28.46154
	422	0	0
	2684	.0041	48
	777.753	.0014284	23.4436
Home Healthcare	1464.818	.0002545	26.85714
	248	0	0
	2640	.0014	47
	703.4584	.000439	25.12256
Pharmaceutical M	1765.407	.0005259	30
	332	0	0
	4392	.0027	48
	875.9515	.0007573	23.45044
Total	2265.139	.0009473	26.64706
	248	0	0
	21573	.0177	48
	2643.873	.0020824	22.63104

Figure 2: Commercial Banking Industry, Big Data as a Percent of the LHS for 2011-2014

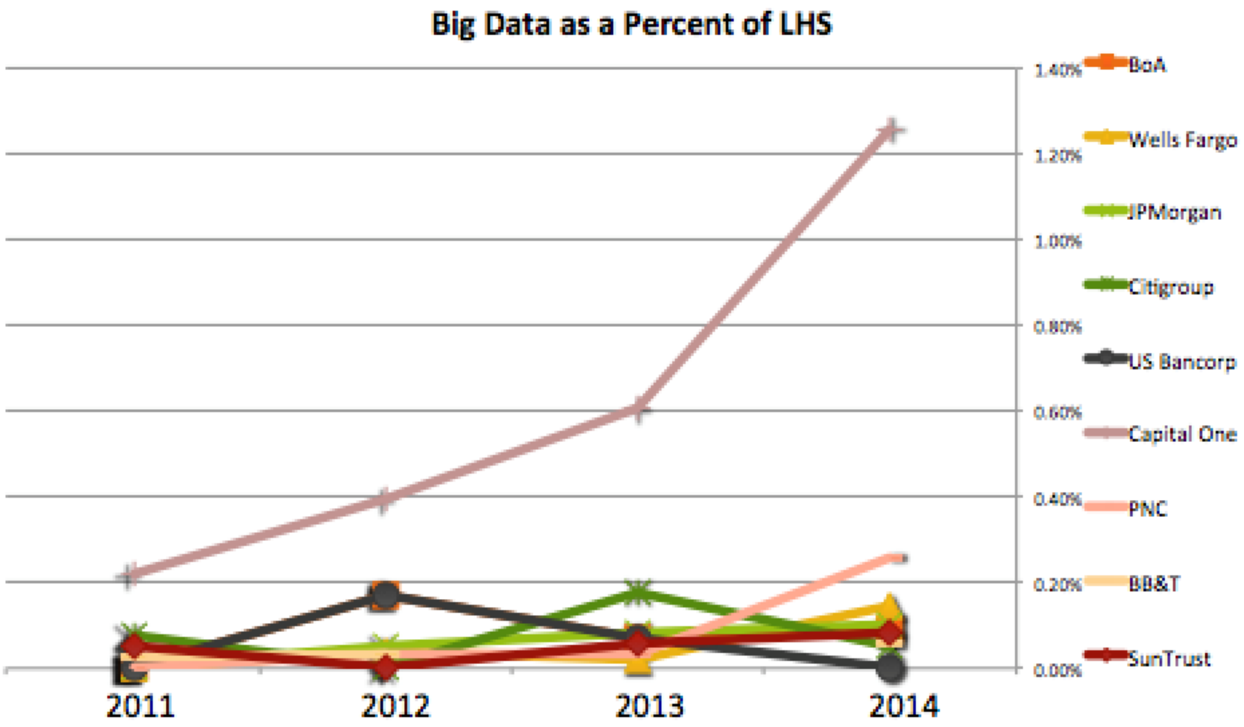

Figure 3: Capital One Compared on Relevant Metrics to Closest Competitors

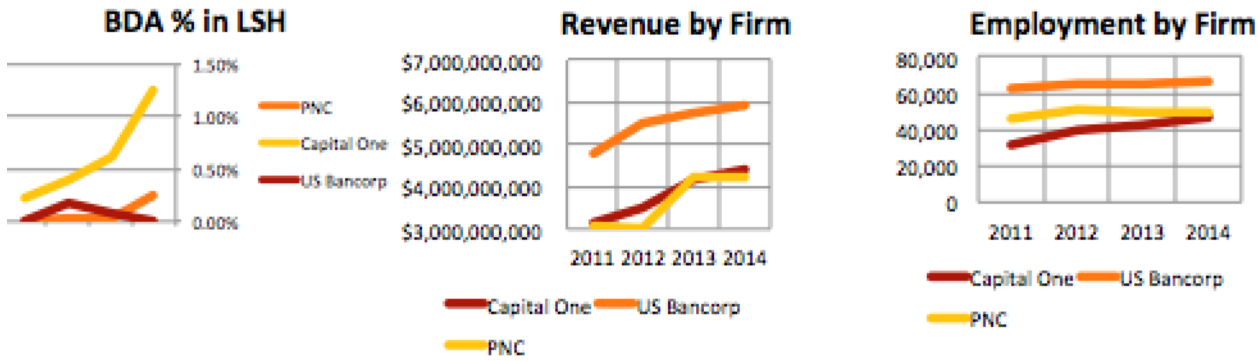

Figure 4: Crude Oil Pipeline Transportation Variable Analysis, Descending by Net Income

Company:	Number of Words	% of LHS	C-Suite Position 1	C-Suite Position 2	Net Income	Number of Employees	Industry	Year
Enterprise Products Partners	2953	0.00%	46	0	$2,833,500,000	N/A	Crude Oil Transportation	2014
Kinder Morgan	979	0.00%	48	0	$2,692,000,000	11,075	Crude Oil Transportation	2013
Enterprise Products Partners	2079	0.00%	0	0	$2,607,100,000	N/A	Crude Oil Transportation	2013
Enterprise Products Partners	3216	0.00%	0	0	$2,088,300,000	6,900	Crude Oil Transportation	2011
Energy Transfer Equity	634	0.00%	0	0	$1,274,000,000	14,433	Crude Oil Transportation	2012
Magellan	1347	0.00%	0	0	$839,519,000	1,565	Crude Oil Transportation	2014
Kinder Morgan	1079	0.00%	47	0	$660,000,000	8,120	Crude Oil Transportation	2011
Magellan	1363	0.00%	0	0	$582,237,000	1,459	Crude Oil Transportation	2013
Energy Transfer Equity	535	0.00%	0	0	$528,247,000	2,477	Crude Oil Transportation	2011
Kinder Morgan	1024	0.00%	47	0	$427,000,000	10,685	Crude Oil Transportation	2012
Buckeye Partners	1663	0.12%	0	0	$274,857,000	1,430	Crude Oil Transportation	2014
Buckeye Partners	2654	0.00%	0	0	$230,551,000	1,020	Crude Oil Transportation	2012
Buckeye Partners	2200	0.00%	0	0	$164,425,000	1,270	Crude Oil Transportation	2013
Buckeye Partners	2373	0.00%	0	0	$114,664,000	1,029	Crude Oil Transportation	2011

Figure 5: Firms Across Industries with CIO/CTO As Second Officer and 0% Big Data LSH

Company:	Number of Words	% of LHS	C-Suite	C-Suite Position 1	C-Suite Position 2	Net Income	Number of Employees	Industry	Year
Merck	332	0.00%	CIO Chris Scalet #2	48	0	$6,392,000,000	86,000	Pharmaceutical Manufacturing	2011
Supervalu	422	0.00%	CIO & EVP Randy Burdick #2	48	0	$192,000,000	38,500	Grocer	2014
Merck	455	0.00%	CIO Chris Scalet #2	48	0	$4,517,000,000	76,000	Pharmaceutical Manufacturing	2013
Merck	696	0.00%	CIO Clark Golestani #2, CIO Chris Scalet #3	48	47	$11,934,000,000	70,000	Pharmaceutical Manufacturing	2014
Supervalu	956	0.00%	CIO & EVP Randy Burdick #2	48	0	$182,000,000	35,800	Grocer	2013
Kinder Morgan	979	0.00%	CIO Henry Neymann #2	48	0	$2,692,000,000	11,075	Crude Oil Transportation	2013
SunTrust	1293	0.00%	CIO Anil Cheruyin #2	48	0	$1,958,000,000	26,778	Commercial Banking	2012
Merck	1351	0.00%	CIO Chris Scalet #2	48	0	$6,299,000,000	83,000	Pharmaceutical Manufacturing	2012
AFLAC	1397	0.00%	CIO Michael Boyle #2	48	0	$1,964,000,000	8,562	Direct Life Insurance	2011
Bristol-Myers Squabb	1474	0.00%	#2 Paul Van Autenreid CIO	48	0	$5,260,000,000	27,000	Pharmaceutical Manufacturing	2011
Bristol-Myers Squabb	1631	0.00%	#2 Paul Van Autenreid CIO	48	0	$2,501,000,000	28,000	Pharmaceutical Manufacturing	2012
Gap	1702	0.00%	CIO Tom Keiser #2	48	0	$833,000,000	132,000	Clothing Retail	2011
J&J	1906	0.00%	CIO Laverne H. Council #2	48	0	$10,514,000,000	127,600	Pharmaceutical Manufacturing	2012

Figure 6: Table of Top 10 Exhibiting Firms by Percent of the LSH Dedicated to BDA

Company:	Number of Words	% of LHS	C-Suite	C-Suite Position 1	C-Suite Position 2	Net Income	Number of Employees	Industry	Year
Razorfish	565	1.77%	CTO Ray Velez #4	46	0	N/A	2,200	Advertising Agency	2011
Capital One	3908	1.25%	CIO Robert Alexander #2	48	0	$4,428,000,000	46,000	Commercial Banking	2014
Ogilvy & Mather	468	0.85%	Worldwide CDO Brandon Berger #9, CDO Todd Cullen #10	41	40	N/A	11,000	Advertising Agency	2013
Capital One	3796	0.61%	CIO Robert Alexander #2	48	0	$4,159,000,000	41,951	Commercial Banking	2013
Gap	1063	0.47%	CIO Tom Kaiser #3	47	0	$1,280,000,000	137,000	Clothing Retail	2013
Ogilvy & Mather	636	0.47%	Worldwide CDO Brandon Berger #5	45	0	N/A	11,000	Advertising Agency	2011
Ogilvy & Mather	446	0.45%	Worldwide CDO Brandon Berger #11, CDO Todd Cullen #12	39	38	N/A	11,000	Advertising Agency	2014
Gap	1411	0.43%	CIO Tom Kaiser #2	48	0	$1,135,000,000	136,000	Clothing Retail	2012
Krogers	2684	0.41%	N/M	0	0	$1,747,000,000	400,000	Grocer	2014
Capital One	4310	0.39%	CIO Robert Alexander #2	48	0	$3,517,000,000	39,593	Commercial Banking	2012

Figure 7: Industry Frequency from 2011 to 2014

Industry	Freq.	Percent	Cum
Advertising Agency	4	2.42	2.42
Clothing Retail	26	15.76	18.18
Commercial Banking	38	23.03	41.21
Crude Oil Transportation	19	11.52	52.73
Direct Life Insurance	27	16.36	69.09
Grocer	13	7.88	76.97
Home Healthcare	11	6.67	83.64
Pharmaceutical Manufacturing	27	16.36	100.00
Total	165	100.00	

Figure 8: Regression When C-Suite Position 0 Were Removed

```
          ofLHS |      Coef.   Std. Err.      t    P>|t|     [95% Conf.
Interval]
----------------+------------------------------------------------------------
----
CSuitePosition1 |   3.66e-06   .0000425     0.09   0.931    -.0000807
.0000881
          _cons |   .0010552    .001923     0.55   0.585    -.0027659
.0048762
-----------------------------------------------------------------------------
```

Figure 9: Multivariate Regression for Correlation between C-Suite and percent LSH

(Std. Err. adjusted for **42** clusters in company)

oflhs	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
number of wo~s	**1.28e-08**	**4.86e-08**	**0.26**	**0.794**	**-8.53e-08**	**1.11e-07**
csuiteposi~1	**6.57e-06**	**6.90e-06**	**0.95**	**0.346**	**-7.36e-06**	**.0000205**
year	**.0002328**	**.0001228**	**1.90**	**0.065**	**-.0000151**	**.0004808**
industry1	**(dropped)**					
industry3	**-.0004954**	**.000796**	**-0.62**	**0.537**	**-.002103**	**.0011122**
industry4	**-.0014445**	**.0004941**	**-2.92**	**0.006**	**-.0024424**	**-.0004466**
industry5	**-.0016938**	**.0004601**	**-3.68**	**0.001**	**-.0026231**	**-.0007645**
industry6	**-.0011975**	**.0005738**	**-2.09**	**0.043**	**-.0023564**	**-.0000386**
industry7	**-.0013622**	**.0005468**	**-2.49**	**0.017**	**-.0024665**	**-.000258**
industry8	**-.0010049**	**.0004852**	**-2.07**	**0.045**	**-.0019847**	**-.000025**
netincome	**-5.56e-14**	**4.03e-14**	**-1.38**	**0.175**	**-1.37e-13**	**2.58e-14**
number of em~s	**2.86e-09**	**2.96e-09**	**0.96**	**0.340**	**-3.13e-09**	**8.84e-09**
_cons	**-.466995**	**.2470579**	**-1.89**	**0.066**	**-.9659386**	**.0319486**

Figure 10: Regression for C-Suite on percent LSH with Crude, Direct and Grocer Removed

```
------------------------------------------------------------------------------
----
          ofLHS |      Coef.   Std. Err.      t    P>|t|     [95% Conf.
Interval]
----------------+-------------------------------------------------------------
----
CSuitePosition1 |   .0000228   .0000103     2.20   0.030     2.26e-06
.0000433
          _cons |   .0007428    .000356     2.09   0.039     .0000367
.0014488
```

Figure 11: Leader Vs. Laggard Dynamic

For leader

reg ofLHS CSuitePosition1 if lead==1

```
     Source |       SS       df       MS              Number of obs =      13
-------------+------------------------------           F(  1,    11) =    0.88
       Model |  .000023307     1  .000023307           Prob > F      =  0.3672
    Residual |   .00028988    11  .000026353           R-squared     =  0.0744
-------------+------------------------------           Adj R-squared = -0.0097
       Total |  .000313187    12  .000026099           Root MSE      =  .00513

------------------------------------------------------------------------------
         ofLHS |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------+----------------------------------------------------------------
CSuitePosition1 |   .0000637   .0000678     0.94   0.367    -.0000854    .0002128
         _cons |   .0011416   .0025547     0.45   0.664    -.0044812    .0067645
------------------------------------------------------------------------------
```

For laggard

reg ofLHS CSuitePosition1 if lead==0

```
     Source |       SS       df       MS              Number of obs =     152
-------------+------------------------------           F(  1,   150) =    1.72
       Model |  3.7547e-06     1  3.7547e-06           Prob > F      =  0.1917
    Residual |  .000327494   150  2.1833e-06           R-squared     =  0.0113
-------------+------------------------------           Adj R-squared =  0.0047
       Total |  .000331248   151  2.1937e-06           Root MSE      =  .00148

------------------------------------------------------------------------------
         ofLHS |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
---------------+----------------------------------------------------------------
CSuitePosition1 |   6.88e-06   5.24e-06     1.31   0.192    -3.48e-06    .0000172
         _cons |   .0005961   .0001743     3.42   0.001     .0002516    .0009406
------------------------------------------------------------------------------
```

References

Accenture. 2014. Big Success with Big Data. *Accenture Data Survey*.
ACM. 2014. Visualizations Make Big Data Meaningful. *Communications of the ACM 57.* (6).
Agneeswaran, V. 2012. Big-Data – Theoretical, Engineering and Analytics Perspective. *Innovation Labs, Impetus Infotech (India) Pvt Ltd.* 1, 13-14.
Aileen, A. and J. Cummings. 2016. The Rise of the Chief Information Security Officer. *People & Strategy 39*, (1).
Andrews, A.R., G. Cheatham, A. Immaneni and A. McCombs. 2007. Capital One Banks on Six Sigma for Strategy Execution and Culture Transformation. *Global Business and Organizational Excellence*, 43-56.
Assuncao, M., S. Bianchi, R. Buyya, R. Calheiros and M.A.S. Netto. 2015. Big Data Computing and Clouds: Trends and Future Directions. *Journal of Parallel and Distributed Computing*, 3-15.
Barnes, D. 2012. Report: Data analytics - Computing infrastructure - Big data becomes big business. *The Banker*.
Bartel, H and D. Scheibenreif. 2014.Business Moment: A Supermarket Mobile App Enhances User Experience and Helps Find New Friends in a New City. *Gartner*.
Behrend, T., M. Karim and K. Willford. 2015. Big Data, Little Individual: Considering the Human Side of Big Data. *Industrial & Organizational Psychology 8*, (4).
Betsy, S. 1999. Multimedia Market Research. *Marketing Research 11*.
Bradley, T., V. Ng and P. Thibodeau. 2014. The Internet of Things – Threats and Challenges. *Network World Asia 11*, (1): 16.
Brynjolfsson, E., and K. McElheran. 2015. Data in Action: Data-Driven Decision Making in U.S. Manufacturing. *CES 16-06*, 3-7.
Burns, M., J. Cao and B. Friedman. 2014. Enterprise Social networking Data Analytics Within Alcatel-Lucent. *Bell Labs Technical Journal 18*, (4): 89-109.
Burt, M., J. Davison, R. Hetu and K. Marian. 2014. Predicts 2015: Retail Digital Commerce Stays Low, Crowdstorming Leads Innovation, In-Moment Pricing Drives Personalization, and Augmented Reality Rallies. *Gartner*.
Chaudhuri, A. 2009. The Classical Continuity of Electronic Globalization. *Decision 36*. (1): 33.
CIO Review. 2015. Big Data Special. *CIO Review*. 10-58.
Cribbs, J. and T. Handler. 2015. Hype Cycle for Consumer Engagement with Healthcare and Wellness. *Gartner*.
Cruz, G., S. Ginel and C. Ordaz. 2015. The Influence of Top Management Team Conflict on Firm Innovativeness. *Group Decision & Negotiation 24*.
Davenport, T., and J. Dyche. 2013. Big Data in Big Companies. *SAS Institute*, 2-30.
Deloitte. 2015. Analytics Trends 2015, Looking Below the Surface. *Analytics Trends 2015*.
De Prato, G., and J.P. Simon. 2015. The next wave: "big data"?. *Communications & Strategies 97*, 15-39.
De Simoni, G., S. Judah and E. Zaidi. 2015. Market Guide for Metadata Management Solutions. *Gartner*.
Duncan, A. D. and D. Laney. 2015. How to Adopt Open Data for Business Data and Analytics

— And Why You Should. *Gartner*.
Duncan, A., A. Linden, H. Koehler-Kruener and E. Zaidi. 2015. Market Guide for Text Analytics. *Gartner*.
Eldred, M. and R. McAvey. 2014. A Compelling Use Case for Server-Based, GPU-Enabled Technology for Oil and Gas Upstream Modeling Suites. *Gartner*.
Fairbank, R. 1997. The Technology Bank. *Chief Executive*, (124): 12-13.
Filbeck, G. and S. Visscher. 2003. Dividend-Yield Strategies in the Canadian Stock Market. *Financial Analysts Journal 59*.
Fisher, S. and M. Hu. 1988. Does the CEO's Letter to the Shareholders Have Predictive Value. *Business Forum 14*, (1): 22.
FTC Report. 2016. Big Data a Tool for Inclusion or Exclusion?. *Federal Trade Commission*.
Geppert, J. and J. Lawrence. 2008. Predicting Firm Reputation Through Content Analysis of Shareholders' Letter. *Corporate Reputation Review 11*, 285-307.
Gerth, A. and J. Peppard. 2016. The Dynamics of CIO Derailment: How CIOs Come Undone and How to Avoid it. *Business Horizons 59*, (1): 61-70.
Grover, V. S.R. Jeong, W.J. Kettinger and C.C. Lee. 1993. The Chief Information Officer: A Study of Managerial Roles. *Journal of Management Information Systems 10*, (2): 107-130.
Guadalupe, M., H. L. Hongyi and J. Wulf. 2014. Who Lives in the C-Suite? Organizational Structure and the Division of Labor in Top Management. *Institute for Operations Research and the Management Sciences 60*, (4): 824-844.
Harris-Ferrante, K. 2016. Transforming into a Digital Insurance Company Primer for 2016. *Gartner*.
Hirsh, A. and R. Spatding. 2012. Risk-Managed Approach for Routing Petroleum Pipelines: Keystone XL Pipeline, Nebraska. *Environmental Science & Technology 46*, (23): 12754- 12758.
Huang, O. and D. Laney. 2015. How Organizations can best Monetize Customer Data. *Gartner*.
Jarvenpaa, S. 1990. Information Technology and Corporate Strategy: A View from the Top. *Information Systems Research 1*, (4): 352-370.
Jensen, J, and R. Perrons. 2015. Data as an asset: What the oil and gas sector can learn from other industries about "Big Data". *Energy Policy 81*, 117-121.
Kohers, G. 2015. SAP and the Introductory Management Information Systems Course. *Academy of Educational Leadership Journal 19*.
Lambert, M. 2015. Pipeline Monitoring Evolution Under Way. *American Metal Market*.
Latimore, C. 2015. Automation in Mining. *Process & Control Engineering 68*.
Lin, W.T. and W.H. Weng. 2014. Development Trends and Strategy Planning in Big Data Industry. *Contemporary Management Research*, (3): 2013-2014.
McKinsey & Company. 2013. Strategy to Beat the Odds. *Quarterly Report 4*.
Morris, J.G, N. Sun, M. Xie, J Xu and X, Zhu. 2014. iCARE: A Framework for Big Data-Based Banking Customer Analytics. *IBM Journal of Research and Development*.
Moyer, K. 2016. The Economics of Connections for Banking. *Gartner*.
News RX. 2011. Mohr Davidow; Mohr Davidow Partners, Portfolio Company Executives to Speak on Big Data's Impact on the Financial Services and Advertising Industries. *News RX*.
Patel, A. 2015. Network Performance Without Compromising Security. *Network Security 2015*, (1): 1-20.

Rotman, D. 2013. How Technology is Destroying Jobs. *MIT Technology Review Journal*.

Sharma, S. 2016. Expanded Cloud Plumes Hiding Big Data Ecosystem. *Future Generation Computer Systems 59.*

Sterling, B. 1997. The Digital Revolution in Retrospect. *Communication of the ACM 40*, (2): 59-ff.

Venkatachalam, N. 2014. Using Big-Data Analytics to Manage Data Deluge and Unlock Real-Time Business Insights. *The Journal of Equipment Lease Financing 80*, 48.

Wade, J. 2012. PART 3: How the Insurance Industry Is Using Big Data to Better Define Risk Appetite. *Risk Management*.