Gordon Institute of Business Science
University of Pretoria

Creating value from big data and analytics: a leader's perspective

Melissa Anine Stevens

16391994

A research project submitted to the Gordon Institute of Business Science, University of Pretoria, in partial fulfilment of the requirements for the degree of Master of Business Administration.

6 November 2017

Abstract

The world is increasingly using technologies that generate and consume unimaginably large quantities of data, called big data. The power of big data does not lie only in its quantum, but in what organisations do with it. Within a business context, big data and analytic methodologies offer the potential to generate unique insights. However, the reality is that many organisations have not yet mastered the art of using big data and analytics to create value.

The objective of this research was to assist organisations that are on the journey to becoming data-led, by exploring a leaders' perspective on the required building blocks of the process through which big data and analytics create value.

This topic was explored through nine qualitative, semi-structured interviews with leaders of financial service organisations operating in South Africa. The study found that organisations that created value from big data and analytics needed leadership support to be able to successfully create a data-led decision making environment. Furthermore, organisations needed diverse skills embodied in staff that were willing to learn continuously, had strong quantitative abilities and business acumen. Different physical infrastructure is also needed, and this created a need for financing. Importantly, organisations also needed to have an understanding of what value they were pursing through a big data initiative.

Keywords

Big data, analytics, evidence-based, decision making, resources, value, financial services.

Declaration

I declare that this research project is my own work. It is submitted in partial fulfilment of the requirements of Master of Business Administration at the Gordon Institute of Business Science, University of Pretoria. It has not been submitted before for any degree or examination in any other University. I further declare that I have obtained the necessary authorisation and consent to carry out this research.

Melissa Anine Stevens

6 November 17

Date

List of figures

List of tables

Table of contents

"Horizontally integrated financial services players have got a significant data advantage in the new economy, in the fourth industrial revolution and they are actually the beneficiaries of the ability to basically have electronic interactions with their customers, using the data to take better decisions …" RP5

Chapter 1: Introduction

Organisations face the challenge of remaining competitive in a fast-paced, global landscape in which disruptive new competitors emerge from unexpected fronts, competitive advantages are rapidly gained and lost, and innovative new technologies are frequently introduced. Disruption, enabled by innovative technological advances, is forcing organisations to rethink who their customers are today, who they will be in the future, and how organisations will continue to create value for them. There is little doubt that organisations need to reinvent aspects of their business.

The pervasive use of the internet has fuelled rapid growth in e-commerce and organisations operating in the digital economy generate vast, ubiquitous data sets, commonly referred to as big data. This big data can be used to create value for organisations. DalleMulle and Davenport (2017) noted that data has gone from being critical to only a few functions within a few organisations, to being pivotal to any business. Furthermore, Wamba et al. (2016) believed that big data and analytics would have a significant impact on a range of industries – from healthcare, to manufacturing. It is clear that the days where only technology firms can create products and services from analyses of big data are long gone: all firms across all industries must leverage big data (Davenport, 2013a). Data is a strategic asset and it should be managed accordingly.

McAfee and Brynjolfsson (2012) emphasised the power of big data and analytics to guide actions taken by leaders. They stated that big data and analytics allowed "more accurate predictions, better decisions, and precise interventions, and can enable these things at a seemingly limitless scale". They noted that many companies that characterised themselves as data-driven demonstrated stronger financial performance. Barton and Court (2012) supported this notion by asserting that using data as a key element to enable strategies could secure a competitive advantage. Despite this, many organisations, and therefore leaders, were not embracing data-driven decision making (McAfee & Brynjolfsson, 2012). Therefore, leaders that have not yet harnessed the ability to leverage big data and analytics to their organisations' competitive advantage faced significant challenges. To remain competitive, such laggard organisations will have to start applying on-going attempts to make sense of big data through the use of analytics (Davenport & Kudyba, 2016).

Based on research, the inability of organisations to create competitive advantage through big data did not mean that companies were not acknowledging its utility or were refraining from investing in big data. In fact, in some cases, the contrary applied. According to Wang (2016) much progress is to be made with big data, as a mere 8% of organisations that had invested in big data and analytic capabilities were doing anything significant with their investments. Most organisations were using big data to generate incremental advances, rather than to achieve the transformation required to gain competitive ground. Henke, Bughin, and Chui (2016) found that the majority of industries were not close to tapping into the full potential of big data and analytics to create further revenues and efficiencies. They noted that the lack of value creation was not due to a lack of investment, and that the 500 business leaders that were surveyed were "only somewhat" using the investment in big data and analytics to achieve the objectives set out when making the investment. This could be indicative that challenges exist in using big data successfully.

On the other hand, Ross, Beath and Quaadgras (2013) offered a different view for why big data was creating so little value. They argued that the benefits and advantages of big data were overhyped. They were not alone in their view that big data was perhaps somewhat of a fad. Corte-Real, Oliveira and Ruivo (2016) posited that big data and analytics was not a silver bullet to creating competitive advantage, and found that firms were not able to create business value from their data and analytic capabilities. DalleMule and Davenport (2017) found that companies were not using their data resources to make decisions, with less than 50% of structured and less than 1% of unstructured data used to inform business decisions. Ross et al. (2013) posited that most organisations are not yet able to use traditional data effectively and hence, despite huge investments in big data and analytic capabilities, they were not able to create value from big data. Ross et al*.* (2013) were not necessarily saying that big data and analytics could not create value, but rather that companies should first start to use the existing volumes of data and experiment with analytics to ensure the organisation is positioned to exploit big data and analytics rather than wishfully using big data as an opportunity to try and overcome existing gaps.

Henke et al. (2016) offered a slightly different take on why value was not created, and noted that an implementation gap exists between experimenting with data and analytics, and implementing the changes in the operations of the business to exploit the insights gained. Therefore, they were alluding to a problem with scaling and integrating data an analytics into the operations. Corte-Real *et al.* (2016) emphasised a need to understand when, why and how big data and analytics can be relevant in creating competitive advantage. DalleMule and Davenport (2017) note that inadequate management of information asset results in an inability to leverage data and analytics, and note that data management is the responsibility of leaders. David Grant (personal communication, 21 April 2017) commented that obtaining information that leaders require to provide more bespoke products and services to consumers presents a significant, unaddressed challenge in practice.

Therefore, as the aforementioned literature illustrates, there is generally a conceptual appreciation for the need to implement big data and analytics to remain successful in the modern economy, however big data and analytics is not an automatic solution to creating competitive advantage. It cannot simply be layered on top of current processes – organisations face a challenge in changing the established paradigms and operationalising big data and analytics to actually create value from big data and analytics.

Despite anecdotal evidence that challenges existed in creating value from big data and analytics, change is imperative as the current environment raises challenges for organisations and leaders employing traditional business concepts (George, Haas & Pentland, 2014). These challenges may be existential. It is evident that leaders have not yet cracked the successful implementation of big data and analytics, and that they perceive tremendous challenges in becoming big data-enabled. Many leaders may in fact be uncertain of how to proceed to exploit big data and analytics as they question the maturity of their organisations. This may result in leaders being wary of making significant investments to position their organisations to take advantage of big data and analytics (Barton & Court, 2012). This was precisely the incorrect response – despite the challenges, it was imperative that organisations persevere through the pain of transitioning to a data driven mode of operation (Henke et al., 2016). Leaders need guidance to help them overcome the challenges. Part of the challenge faced by leaders is knowing what the right questions are to ask, and what data analytics are required to appropriately identify and meet their consumer's needs.

Turning to academic sources for guidance on what the challenges are and how they could be overcome also may not provide leaders with much useful information, especially as it relates to the challenges around meaningful analytics. Wamba, et al. (2016) noted that big data is a hot topic amongst scholars and that mixed views exist about its value in creating potential. Recent studies acknowledged that business analytics was a challenging new scope extension of operational research, with surveys of practice indicating that business analytics embodied changing practitioner perspectives and activities, yet academic journals continued to include comparatively low studies in this field (Mortensen, Doherty & Robertson, 2014; Ranyard, Fildes & Hu, 2015). George, Osinga, Lavie and Scott (2016) noted that academically, many areas of practice, notably strategy, has not been explored in sufficient detail. Corte-Real et al. (2016) noted that literature indicated that room existed for further research in the use of big data and analytics, especially as it related to post implementation success, and how to leverage these to create competitive advantage. George *et al.* (2016) noted that big data and analytics was a new, evolving field that presented untapped opportunities for scholars and practitioners alike, which could lead to new questions in current and new areas, and better answers in response to these. There was a call for combining big data with qualitative techniques to add greater depth and more meaning to big data as it provides context to it (Boyd & Crawford, 2013; Crawford, 2011).

Kiron and Shockley (2011) and Vidgen, Shaw and Grant (2017) explored the challenges that management faced in using business analytics, and provided recommendations for overcoming these. However, further work was required, particularly around the challenges leaders face as they guide organisations striving to using analytics to provide increasingly more bespoke goods and service offerings to their customers. Understanding these processes will enable a better assessment of whether the substantial investment in big data and analytic capabilities should be undertaken, as leaders can get a clear understanding of any gains which may arise from successful implementation of big data and analytics. It is to this body of literature that this study aims to contribute.

1.1 Problem statement

The objective of this research was to understand the leader's perspective of the practicalities involved in creating value from big data and analytics.

1.2 Significance of this study

This study aimed to be beneficial to leaders of organisations that need to transform to enable them to remain or become competitive in the modern digital economy. By understanding key elements of the process through which organisations have become data-led, the leaders of organisations that are embarking on a transition to being data-driven will be able to gauge organisational readiness and better anticipate challenges in transformation and, potentially, the means to overcome these. This should lead to less set-backs in organisations' ability to produce products, services and features that are valued by consumers.

The next section of this study offers a reviewing of the available literature in more detail to identify sources that could shed light on the process for creating value from big data and analytics, including the challenges faced, and means of overcoming these.

Chapter 2: Literature review

2.1 Introduction

The objective of this research was to determine the process through which value can be created from big data and analytics, from the perspective of leaders. McAfee and Brynjolfsson (2012) noted that there was great academic interest in the supposed value of being data-driven, as evidenced by the rifeness of anecdotes and case studies on this topic in the business press. Yet, this view was not held by all scholars. George *et al*. (2014) called on ‘’management scholars to unpack how ubiquitous data can generate new sources of value” and indicated a shortage of academic interest and material on big data research. These contradictory views were symptomatic of some of the contentiousness perceived in big data and analytics research as a management and leadership construct, which this literature review has explored.

This literature review has explored academic literature on the concepts of big data, and analytics as the most obvious areas of relevance to the objective of understanding how value can be created. To contextualise the value creation process, the literature has also explored the application of big data and analytics in decision-making and the implications for, and observations by leadership in the process of creating value from value through the use of big data and analytics. To be able to make sense of the objective of using big data and analytics, the concept of what could be considered to be value is also explored. At various points in the literature review that follows, text boxes are included containing key thoughts that are carried forward into a conceptual model that is presented at the end of this chapter.

2.2 Big data

The terms “big data” and “big data analytics” have been used to describe the large, complex data sets and analytical techniques applied to search through data to find correlations and ultimately derive business insights (Davenport, 2014). However, big data and analytics are separate, although complementary topics, and hence the literature review addresses it under distinct headings starting off with this discussion of big data.

The use of data to make decisions is not a new idea. Small amounts of data have historically been used to support internal business decisions (Davenport, 2014). However, many organisational leaders still apply a fair amount of intuitive decision making, especially where the data appears to contrast with intuition (McAfee & Brynjolfsson, 2012). There is consensus that the recent advent of big data has significantly changed the nature and scope of data-enabled decision making to create and capture value for business.

Big data differs from traditional data sources and systems in that it is characterised by the three Vs:

1. *Volume:* Big data represents a vastly increased and ever increasing quantity of data. Due to cost reductions created by advances in data storage and mining, companies now have the opportunity to work with unimaginably large data sets. Seddon and Currie (2017) therefore emphasised that "volume" refers to the capability to manage the data by being able to effectively store and retrieve the data. For example, they noted that on an hourly basis Wal-Mart collected more than 2.5 petabytes of transactional data which would have required 20-million filing cabinets were it not for technological advances (Vidgen *et al.*, 2017). Increasingly, it is acknowledged that it is not just the quantum of data that is an important characteristic of big data, but rather it is the insights that the data can provide. In this regard, the granularity of the big data provides vastly superior contributory information as it enables an understanding of not only the particular outcomes achieved, but also the nuances attached to those outcomes (George *et al.*, 2014).
2. *Velocity:* Big data is generated at greater speed compared to traditional data. For example, mobile phone location data has been used to estimate Macy's sales, even before the actual sales were recorded (McAfee & Brynjolfsson, 2012). Real-time or near real-time information provide rapid insights that can provide organisations with a strong competitive advantage. Seddon and Currie (2017) noted that velocity was not only about how quickly big data was gathered but also how quickly it was processed.
3. *Variety:* "Variety" alludes to the varying nature of big data. Big data is comprised of multiple types of data: unidimensional or multidimensional, as well as structured or unstructured. Examples of data include text, images and video – or even a combination of these. Another aspect of "variety" is that big data also finds its origin in various sources, such as location signals from mobile devices, business and purchase transactions, social media, and data from sensor networks (McAfee & Brynjolfsson, 2012). While many organisations have worked extensively with structured data, the challenge in commercialising big data lies in the unstructured data – which comprises the biggest new opportunity in the big data world. Within financial services, the interpretation of unstructured data, such as that derived from social media, provides the newest challenge, while voice and video data are similarly nascent (Seddon & Currie, 2017). Seddon and Currie (2017) noted that new programming languages such as Julia, open source software, and storage infrastructure such as Hadoop, are being developed to cater for the transformation of unstructured data into data that can be quantitatively processed by computers. The advent of the new programming and infrastructure norms are indicative of the new capabilities which may be required by organisations to be able to create value from big data.

The three Vs discussed above not only characterise big data, they are the key source of the power of big data. A number of publications also refer to the five, six or even seven Vs of big data, with veracity, value, variability, and viability referenced as additional characteristics of big data (Gil & Song, 2016; Biehn, 2013). Seddon and Currie (2017) refer to the seven Vs by attributing veracity, value, variability and visualisation as big data characteristics. They note that visualisation is the ability to present the trends and patterns in the data. Value refers to a firm's ability to compete based on data and analytics output to return a profit. Variability refers to the changing insights gained from various interpretations of the same data or data that is augmented from different sources.

Lugmayr, Stockleben, Scheib and Mailaparampil (2017) argued that the Vs view of big data is obsolete and instead proposed that a new concept, "cognitive big data", was required. This definition will characterise and create an understanding of big data as a knowledge processing phenomenon which aids humans in their cognitive endeavours, such as their decision making. The additional Vs plus Lugmayer et al. (2017)'s alternate view of how big data can be defined is indicative of the fact that big data as a topic is still subject to evolution. Arguably, this adds a dimension of challenge to organisations that who to harness this vast, complex resource around which conceptual thinking is still ongoing.

Big data is also different from traditional data in how it is aggregated and managed. Not only internal, but also external data sources may be combined to create valuable big data sources (Willmott & Dewhurst, 2014). An argument is also made for a change in the approach of data governance through the democratisation of data within organisations, centralising that which is required to be kept private due to regulation, for example, while decentralising the storage of that which can be shared freely across an organisation to ensure ease of access to data – actually encouraging the wide use of big data (Saran, 2015). Ross et al. (2013) added that the benefits of the information economy will be used when all people are able to proficiently use data. Data management practices have evolved to accommodate and enable the new management approaches required for big data, for example data lakes which have been developed with practically endless capacity for storage of structure and unstructured data that can be democratised, to supplement the use of more high security, controlled data warehouses (DalleMulle & Davenport, 2017).

Davenport (2014) noted that big data is a fast flowing deluge of data and requires ongoing analysis and action, in contrast with a more periodic, ad-hoc and static use of traditional data. However, it was critical that the information was not simply produced and viewed by leaders, but also used to drive decision making. He argued that processes should be defined to aid in managing the insights gained from big data – i.e., when decisions or actions are necessary. In this way a leader would not be bombarded with information without being able to determine which insights, new products, or services should be actioned immediately; what was interesting but not worthy of action; and what was to be implemented later – perhaps in combination with further insights that are yet to emerge. In this way leaders could be more accountable to create valuable output due to having a lack of clarity in their process for placing ideas into production.

It is clear that a big data opportunity exists if an organisation has an enormous amount of data that is not stored in an ordered format such as rows or columns, but is nevertheless crucial to the business as a source of insights through the use of an integrated mix of analytical procedures (Davenport & Patil, 2012). What is also apparent is that there are wider financial, infrastructure, skills and managerial implications and changes required within the organisation to accommodate the effective use of big data.

Key thoughts: Organisations need different finance, infrastructure and skills to create value from big data and analytics

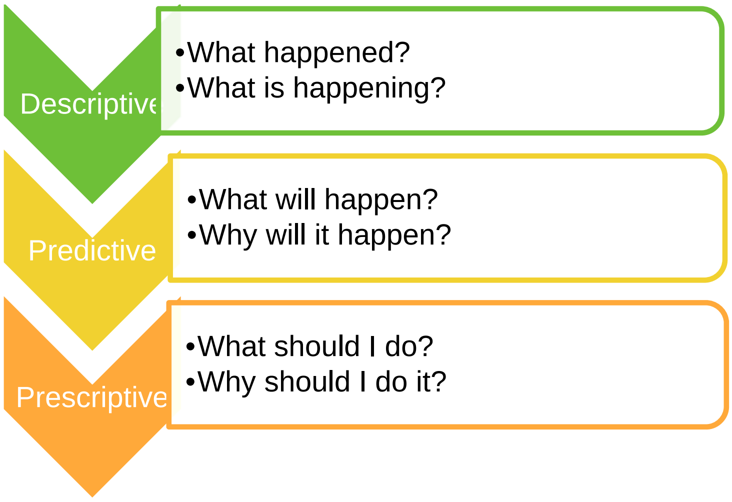
2.3 Analytics

It is true that big data is still just data. To become information, it needs to be organised, aggregated or segmented, analysed, visualised and otherwise manipulated. Analytics are a range of specialised, powerful techniques that could produce information, discern patterns, and predict outcomes more quickly and accurately than the unassisted human mind (George et al., 2014; Davenport, 2013a). Delen and Demirkan (2013) take a more output based view and note that analytics is the processing of analysing trends, building predictive models, and optimising business processes to maximise business performance. Bringing these views together, data analytics has been described succinctly by Corte-Real et al. (2016, p.9) as "a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data, by enabling high velocity capture, discovery and analysis". It enables an organisation to discover information and formulate actionable knowledge about their customers, products and services and evolving market environment. Analytics are increasingly in demand due to the fact that it provides decision makers with much needed accurate information and insights in a timely manner, allowing them to effectively cope with the increases in the complexity of doing business (Delen & Demirkan, 2013).

The effectiveness of analytics depends on the quality and quantity of data, the integrity of its management, and the sophistication of analytical tools used to process the data (Delen & Demirkan, 2013). Due to the enormous volume of data and its complexity, big data may result in spurious correlations and provide information that is not useful to leaders in decision making if it is not processed appropriately. It could also be easy to identify causation where only correlation exists. This highlights the importance of using the right data together with appropriate analytic methodologies to provide decision-useful information. Boyd and Crawford (2011) also noted that large data sets come with inherent issues which are often not acknowledged, making it appear that bigger data sets are always better, when this is not necessarily the case. Instead they argue that the data size should be appropriate to enable the researcher to answer the question posed. The quality of data is important as large data sets often contain weaknesses, such as errors and gaps and sampling biases which are exaceRP2ated by combining various data sets, hence it is important to understand the providence of data.

Mortensen et al. (2014) noted that the volume of data presents a significant challenge for both technology, and the quantitative methods used to process data. This means that as big data has emerged and evolved over time, analytics have also had to be adapted. Davenport (2014) cited the fact that in the past, that traditional data analytics has been used in internal decision making, for example in creating reports. On the other hand, more recently, big data analytics has been used to drive discovery and experimentation, leading to the creation of high value customer experiences and to competitive intelligence that supports strategic decisions – overall improving agility to support the incremental improvements required to stay relevant in an ever changing world. However, it is not that contemporary analytics have replaced more traditionally applied forms. Kiron and Shockley (2011) opined that "baseline data analytics", which enables better management through, for example, enhanced supply chain management, budgeting and forecasting are widely used. Bell (2015) suggested that the basic analytics capabilities that were essential for survival and no longer offered a presupposed source of competitive advantage, instead a deliberate effort, an evolution, is required to achieve the long term benefits. Delen and Demirkan (2013) offered support and described analytics as evolving from descriptive to predictive and finally prescriptive in nature as different types of analytics have emerged to deal with the increasing complexity of big data and the business questions they bring. This evolution is illustrated in Figure 2.3.1. As illustrated, they viewed descriptive analytics as the base level analytics in which business reporting enables the identification of well-defined business problems and opportunities. At the next level, predictive analytics make use of data mining and similar techniques to aid forecasting with accurate projections of the future. Finally, analytics evolve to become prescriptive and with expert systems offer the ability to simulate and model decisions to enable the identification of the most optimal business decisions that should be undertaken and what their outcomes were.

Figure 2.3.1: Evolving nature of business analytics (Delen & Demirkan, 2013).



Support for the evolutionary nature of analytics is evident in other literature as well. Anecdotal evidence suggests that the leadership in many large organisations are using big data analytics to process large internal data sources to improve their decision making and extract operational and financial information that lead to value creation (Vidgen et al., 2017; McAfee & Brynjolfsson, 2012; Davenport, 2014). This may be likened to the descriptive analytics envisioned by Delen and Demirken (2014). However, it is not this level of analytics that can enable big data to be leveraged to its full capability for sustained competitive advantage. A more advanced form of analytics is required. In agreement with the concept of predictive and prescriptive analytics, Delen and Demirkan (2013), Davenport (2013a) echoed the sentiment that analytic capabilities need to expand, and noted that the true value of big data and powerful analytics is not only in producing information for internal decision making, but to drive the development of valuable customer products and services. Contemporary analytics should aim to go beyond simply producing information from data, but also to turn big data into insights. This level of value has not been widely extracted, and Davenport (2013a) argued that organisations needed to fundamentally rethink how they analyse data to create value for themselves and their customers. Amongst others, Bell (2015) suggested that analytics should be applied to the right problems – the critical, strategic issues organisations were grappling with rather than problems that were easy to solve and yet added limited value.

Big data and analytics can offer invaluable insight and generate a competitive advantage but only if the right technology and resources are brought together, and this requires an understanding of why and how technology is used (Corte-Real et al., 2016).

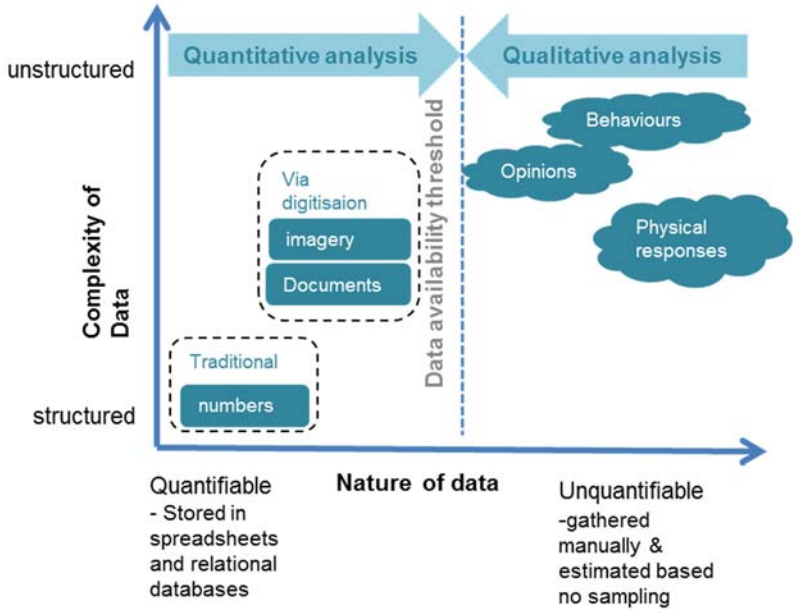
The increased demand for strategic analytics to create deep insights has resulted in a high level of demand for people who poses the requisite skill set to find insights through analytics. Davenport and Patil (2012) referred to such people, who are able to discover new insights while being immersed in data, as data scientists. They are able to structure data and make analysis possible and the output easy to consume, thus enabling decision makers to use data on an ongoing, rather than ad-hoc, basis. These data scientists are also able to participate more fully and unpack the implications of the data output to design new products, processes or formulate a new business direction. In practice, the skill sets of traditional “quants” are perceived as being limited due to their immersion in their technical practice of actuarial science, statistics and mathematics. Their traditionally poor social skills mean that people with analytic capabilities are unable to naturally build sufficient knowledge of the business environment, nor sense and grasp business problems and apply their skills and engage with business to solving these. This prompted Davenport and Patil (2012) to note that data scientists were rare. Data scientists have the ability to bridge into the business world as they have the ability to communicate in a language that is understood by the various stakeholders, and are able to deep-dive into a problem and find the real underlying questions, formulate and test a hypothesis. Ostensibly, to add value, data scientists must be able to place themselves in the position of their customers to anticipate and solve problems.

Delen and Demirkan (2013) also highlighted that analytics require other resources- such as service-oriented, cloud-based resources and infrastructure that permit the necessary capability, scalability yet flexibility to enable the realisation of the value offered by big data and analytics. This is reflective of the fact that data and the required analytic capabilities are not simple plug-in solutions, but requires detailed planning, process overhaul and integration into the operational infrastructure of an organisation to enable it to become agile and effectively deal with the complexity of the current business environment. Given that analytics has been complex, expensive and time-consuming to develop, it was vital to develop appropriate analytics (Davenport & Kudyba, 2016) and to optimise its use in revolutionising business. Barton and Court (2012) suggested that the desired business impact must guide the approach to using big data and analytics, to avoid the wastefulness of mining data without an aim in mind. Willmott and Dewhurst (2014) echoed this and noted that leaders must find methods that enable them to prioritise data sets and insights which will be escalated to their level. Automated decision making needs to be implemented, guided by parameters set by management. They argued that there is an increasing analytical complexity of the world leaders operate, and these leaders need to become comfortable with viewing data and understanding algorithms so as to grasp insights first hand, as opposed to relying on screening and filtering by data scientists or quantitative staff which may dilute the richness of the output. Arguably, in this manner, management would also be able to better understand how algorithms work and therefore grasp and manage the risks arising from their decisions. They also proposed that democratising data and allowing decisions informed by data to be made at lower levels of management is a key enabler of harnessing the power – once again, parameters for what those decisions may entail will have to be set. In this world, the role of leaders will change slightly – the new role will be to leverage their expertise to focus on asking the right questions of the right people at the right time and allow data and analytics to be applied in answering the questions by delivering deeper insights than leaders could achieve on their own. The question remains as to how this can be done.

George et al*.* (2014) and Willmott and Dewhurst (2014) suggested that focussing on data outliers, rather than averages is what leads to critical innovations as this may provide weak signals of trends to come. Willmott and Dewhurst (2014) noted that this experimentation must be led by the CEO, who they refer to as the "chief experimentation officer" (p. 4). Importantly, Willmott and Dewhurst (2014) noted that data outliers could also be the result of problems, rather than bright new opportunities for innovation, which leaders could manage more proactively, potentially improving customer experience and managing the organisational reputation better. The textured, granular nature of big data offers opportunities to identify nascent signals of changing needs. This could be one approach to ensure that big data and analytics were applied to the right questions at the right time.

Irrespective of how analytics are characterised, used, or changes the scope of work of data processors and leaders alike, it is evident that it is grounded in data mining and statistical analysis, and is an ongoing area of research (Chen, Chaing & Storey, 2012). A further opportunity to be explored relates to how the increasing volumes of unquantifiable data, such as sentiment deduced from social media feeds, may be analysed and turned into actionable insights. Unstructured data may be the richest sources of fresh insights which could result in competitive advantage (Willmott & Dewhurst, 2014). Therefore, as the level of quantifiability and structured nature of data decreases, it offers more promise. Despite this, literature on analysing these forms of data, which lends itself to qualitative analysis, is scant in comparison to that related to traditional methodologies for the structured data which has been analysed with increasing sophistication over the last decades. Citi Business Advisors (2015) offered a diagrammatic illustration of this landscape, as per Figure 2.3.1. The y-axis of the diagram below (Figure 2.3.1) illustrates the more extreme spectrums of the data continuum of structured data, such as numbers, to unstructured data, such as human behaviours, as well as non-exhaustive examples of what could lie in between, such as images. The x-axis sets out the extent to which varying data are quantifiable or not- with numbers more quantifiable than human responses. The figure illustrates that data have different characteristics of complexity and different nature and that this may determine the best manner in which it may be analysed- using quantitative or qualitative techniques, or a combination thereof, as discussed in the following subsection.

Figure 2.3.1: Data continuum and the change in analytics methods (Citi Business Advisory Services, 2015).



Key thoughts: Organisations need different decision making, technical skills, leadership skills, processes for asking right questions and infrastructure to create value from big data and analytics

2.3.1 Thick data and ethnography

Boyd and Crawford (2011) viewed the impact of big data as being more profound than its computational ability, as it reframes the concept of knowledge, the research process, and engagement with information. They noted that big data offers an opportunity for gaining insight into human interactions and their immediate society. They countered these favourable statements with a number of cautionaries around big data, and note that there is over-confidence to how big data is viewed, as it is perceived as being able to side-line other forms of research, yet it is not without its weaknesses and limitations. Therefore, despite the enthusiasm around big data and analytics, a school of thought cautioned against the exclusive use of big data as the holy grail of generating insights. This argument was rooted in the fact that it is claimed that the strength of big data is attributable to the quantity of data, but the quality of the insights gained may not be appropriately preserved and valued. Crawford (2013) supported the view that big data has weaknesses, which were underplayed. A key weakness noted was that that big data is subjective, as it is open to interpretation by humans, and therefore vulnerable to human biases (Crawford, 2013). This is because data cannot stand on its own; it requires context. Without understanding the actual context that relates to it, inferences around that context may be drawn, and those inferences are subject to, for example, the character and culture of those using the data. Without an understanding of data's actual, real context, inappropriate decisions may be made. To manage these weaknesses, Crawford (2013) suggested that the data user must understand the origin of the data that had been used, the methods of processing and analysis that were applied to it, as well as their own biases brought into the interpretation of the data. However, context also goes beyond this as it requires an understanding of the subject matter from which the data is drawn- often these are customers or potential customers of an organisation. Despite this, Wang (2016) made the point that big data is currently limited to focus on the capturing and processing of quantitative data and, as a result, human-centred inputs which are descriptive and qualitative in nature are overlooked.

To overcome this, Crawford (2013) stated that big data can be more powerful if it were paired with qualitative approaches to gain a better understanding of the reason and nature of occurrences captured in data, rather than simply its number. One of the methods prescribed to achieve this richer understanding of big data is the application of ethnography. Ethnography is a research strategy that is rooted in anthropology and is concerned with learning from people rather than studying them and hence, it is one method that is able to provide greater contextual understanding (Saunders & Lewis, 2012). Wang (2016) noted that even though the data sets involved in ethnography were often small, it is rich in information unveiled by qualitative, ethnographic research methods and can reveal emotions and paradigms of the world and of the people it relates to, adding depth to what the data can reveal, resulting in the term "thick data". Therefore, the other dimensions of that are data important can emerge.

As discussed, big data is inherently stripped of context and meaning as it needs to be processed, for example standardised and stripped of identifiers. In this process it loses some of its richness – information around the individuals behind the data that could actually be useful to the organisation. Thick data captures these lost elements and creates value as it provides a human-centric layer of information (Wang, 2016).

In essence, big data and thick data produce different types of insights based on the different variables, nature and processes relevant to thick data and big data, as summarised in Table 2.3.1.1. Thick data obtains meaning from the perspective offered by human learning, and is not about quantity but quality, with focus on the depth of understanding of the connections between points. As a result it needs to be tolerant of complexity that cannot be reduced even though it does not have the scale of big data. In contrast, big data is focussed on large quantities, with perspective offered through machine learning. It loses depth, or resolution, as it has to isolate a few variables to process.

Table 2.3.1.1: The difference between thick data and big data (Wang, 2016)

Thick Data	Big Data
Relies on human learning	Relies on machine learning
Reveals the social context of connections between data points	Reveals insights with a particular range of quantified data points
Accepts irreducible complexity	Isolates variables to identify patterns
Loses scales	Loses resolution

Rather than organisations being faced with the uncomfortable prospect of having to make a decision between thick data and big data, an integrated approach may be adopted. In such an approach, big data is used in combination with thick data to create a more complete information set and therefore to aid more robust decision making (Wang, 2016). Although expressed differently, Thompson (2013) shared the view that decision making in a big data world should be not be undertaken based purely on data and algorithms, without pause for deliberation by humans, specifically around the ethical and moral underpinnings of that decision. To understand the moral implications of actions, a richer understanding of the human aspects of the information being considered is required. Therefore, although decisions may be made more quickly and consistently through the use of big data and analytics, its sole use could result in a lack of proper reflection, and this human process remains vitally important to society. It is in this reflection process that thick data becomes relevant.

Occasionally, however, the exclusive use of thick data on its own is appropriate, such as when an organisation is exploring a new venture and an initial understanding needs to be grasped, or when an emotional decision needs to be made (Wang, 2016). Thick data implementation also poses its own challenges: as with data scientists, ethnographers are in short supply (Wang, 2016). Additionally, thick data is typically reported in stories which consume time and resources, and require strong communication skills. Furthermore, the integration of thick data brings to light its own set of business challenges around the best means of reporting its findings to management, the definition of success, and the training of teams to integrate big and thick data approaches. The complexity of applying qualitative means of gaining knowledge to inform decision making is undeniable, but appears to be of significance to successful data-driven business practices. Arguably, to understand the techniques that would best suite a particular business problem, the intended outcomes should be understood.

Key thoughts: Organisations need different decision making processes to create value from big data and analytics

2.4 Value from leveraging big data and analytics

According to Lindgreen, Hingley, Grant and Morgan (2011, p.207), “the creation of value is paramount to any company's survival, especially when dramatic changes lead to fundamental shifts in what companies analyse, create, and deliver’’. Value is a broad term, and yet, for an understanding of the intended outcome of leaders successfully leveraging data, it is important to somehow define or characterise value. An organisation which is looking to start exploiting big data might have a multitude of goals, including improving customer experiences, creating more optimal processes, or enhancing the relevance of their marketing.

To be able to create value, rather than just undertaking endless experiments of questionable value, organisations need to be able to clearly specify how it will be used to generate value and how the results will be measured (Henke et al., 2016). As metrics to gauge results, they noted that a big data initiative can decrease customer acquisition costs by 47% and improve turnover by 8%.

Corte-Real et al. (2016) posited that through the use of big data and analytics, competitive advantage and organisational agility can be achieved by an organisation. However, they noted that organisations did not have clarity on how to create this business value, and companies are essentially proceeding with initiatives without understanding why and how value would be created. Kiron and Shockley (2011) concurred to some extent and explored the generation of competitive advantage as value, and found that 57% of participants in their survey were gaining competitive value from big data and analytics. Davenport (2013a) noted that value as a competitive advantage can only be derived if organisations leverage data and analytics to create products that are more valued by their customers.

McAfee and Brynjolfsson (2012) suggested that factual financial and operational metrics could be used to define value. They found that organisations taking the lead in their industry in using data-driven decision making achieved 5% higher productivity and 6% higher profitability outcomes as compared to their competitors, after controlling for other potential contributors to the results, such as labour. They also cited increased market capitalisation as a viable measure of value created by big data and analytics. Importantly, they noted that early adopters could generate profits faster, entrench their position as market leaders in industries that were slow to adopt big data and analytics, and potentially derail the plans of would-be disruptors of their industries (Henke et al., 2016).

Davenport and Kabyla (2016) opined that large bodies of data and related analytics could in itself represent digital assets, which could be made available to customers and thus create value directly and as a revenue source. They cited Google Inc. and Facebook Inc. as the pioneers of this practice, and ostensibly, it could be argued that more and more companies will be able to engage in similar practices, if they leveraged their own data and analytic capabilities well.

Lindgreen et al. (2011) believed that effective and efficient mobilisation, coordination and deployment of resources were crucial in driving value creation. This means that leaders will only transition to a data-led and analytic-enabled decision making if they were convinced that gearing their resources toward this orientation will enable them to create value. However, they noted that organisations did not have clarity on how to create this business value and essentially companies are proceeding with initiatives without understanding why and how value would be created.

Irrespective of what an organisation may view as value, it is clear that a decision making process will be applied to derive the outcome that will deliver value. The impacts on decision making also need to be understood.

Key thoughts: Organisations need to be able to clearly specify what it sees as value and how it will measure the value it intends to create with big data and analytics

2.5 Decision making

Decision making is a pervasive part of what leaders do and permeates all disciplines and aspects of an organisation, hence it is a key part of exploring the process of leveraging big data an analytics in creating value. Evidence-based management is a family of practices that enable organisations to make decisions based on scientific evidence (Briner, Denyer & Rousseau, 2009; Rousseau & McCarthy, 2007). Part of its definition includes the explicit use of information by leaders to make decisions, and this information can be in the form of experience and judgement, or evidence derived from the environment. Thus, both intuition, which draws on a form of past experiences, and using internal data that may give insight and evidence, forms part of evidence-based management. Historically, management have made decisions by relying more strongly on "intuition", whereas analytics could enable managers to go beyond intuition when making decisions (Davenport, 2013a). Convincing executives who rely heavily on intuition to become more data-driven is challenging (McAfee & Brynjolfsson, 2012). However, for big data and analytics to have value to an organisation, it is clear that it is data-enabled decision making that leaders should conscientiously apply as a less biased, insightful means of making decisions.

Ross et al. (2013) noted that the first step to becoming data-led was to undergo a disruptive culture change to concisely apply evidence-based decision making. Their study of more than 51 organisations found that there were few companies that use data consistently to make decisions, despite the fact that companies that used data-enabled evidence-based decision making were more profitable than their counterparts. They postulated that consistently applying evidence-based decision making requires a culture shift, which is difficult to achieve. The changes required are disruptive as organisational structure, processes, and role definitions also had to change to accommodate the efficient diffusion of big data and analytics and decision making based on it. Only once decision makers have data at their fingertips and use it to make sound, fact-based decisions will organisations be able to use big data to generate operational improvements and profitability that will be hard for competitors to erode. By generating proactive and forward-looking information, big data and analytic capabilities will become a key part of organisational decision making and separate the high and low performing firms (Wamba et al., 2016).

However, as attested by Learmonth (2006), not all practitioners are convinced that evidence-based management remains appropriate in all circumstances as some problems, for example in social studies, are characterised by divergent views and additional evidence does not resolve this divergence. Instead it created more confusion.

In an attempt to bridge the gap between behavioural science and practice, Rosseau and McCarthy (2009) made a case for the incorporation of evidence-based management as an explicit part of the formal education of managers to promulgate the use of evidence to reduce ineffective decisions and to promote substantive expertise and consistency in decision making. The premise for this was that principles should be defined to identify problems which analytics can unmask as generic rather than unique, and therefore instead of repetitively spending an inordinate amount of time on essentially the same issues, it can be solved effectively through the use of big data and analytics. Although their research did not link directly to the use of data in decision making, the foundational elements, such as being able to recognise a decision, and learning the ability to think crucially and to ask appropriate questions to address that decision, is an important linkage into using data and analytics effectively and efficiently.

Saran (2015) quoted Debra Logan, the vice president of Gartner, and noted that the required change in leaders' decision making needed to be that they start by asking what business outcomes, for example, new revenue streams, should occur as a result of a project, such that every information technology (IT) discussion becomes a business outcome discussion, touching on the importance of asking the right question about which problem needs to be solved to get the appropriate answer. Leaders need to educate the organisation overall to be more knowledgeable on what questions they can and should ask. The implication is that organisations need to learn to walk before they run, when it comes to using data in decision making.

Key thoughts: Organisations need different decision making processes to create value from big data and analytics

2.6 Leaders' perspective

Henke et al. (2016) noted that the right capabilities are critical to the successful implementation of big data and analytics. In this regard, they referenced data scientists who have the technical skills to process the big data using meaningful analytics as being important. But, they noted that it is also important to have in-house people with the right business skills and corporate knowledge to ask the right questions to identify business problems and translate the results into business solutions. These people operate as partners with the data scientists in creating value. Specifically, in an organisation where big data and analytics practices are nascent, leaders are these key people. Understanding what challenges these leaders will need to overcome will help inform the skills and capabilities they need to be able to play this important role.

Leadership is a widely studied topic and many definitions and characteristics of leadership exist. Daft (2011, p.5) defined leadership as "an influence relationship which exists among leaders and followers, who intend real changes and outcomes that reflect their shared purpose". Raelin (2015) also noted that leadership is a collective practice observed by people working together, rather than simply being about an individual. The big data and analytics evolution has changed the landscape in which leaders operate and demand new leadership qualities as compared to a few decades ago. Davenport (2013b) maintained that companies need general managers to partner with data scientists. These general managers will play the critical role of building propositions based on the insights produced by the data scientist and combined with the general manager's knowledge of the business. To be able to partner successfully, general managers need to understand analytics, including the fact that it is founded on assumptions. The general manager and data scientist also need to build trust, so that the general manager is able to ask the data scientist tough questions, and to enable the two parties to freely swop ideas to solve business problems, rather than just mathematical problems. Leaders also need to be able to frame a problem as an important first step in the analytical journey, based on their business experience and intuition. General managers also play an important role in relaying the message back to stakeholders in an understandable manner.

Daft (2011) believed that the modern leader should strive to empower rather than to control followers; to encourage collaboration rather than to pit them against each other for the sake of competition; to value differences amongst people; and to revere ethical leadership rather than to act egoistically. These leadership concepts pose challenges for leaders who are called on to operate in greater cooperation. Henke et al. (2016) noted that leaders, including CEOs, need to champion big data and analytics to overcome resistance to its use within the organisation, and to encourage departments to work together in leveraging it. Considered in an environment in which big data and analytics are to be leveraged as a source of value, it implies that a leader would do well to (amongst other things) to empower staff and to collaborate widely to achieve success, which may be challenging, especially in large, multinational organisations.

Vidgen et al. (2017) noted that a point of departure, or overarching challenge, in leveraging big data and analytics is creating a "big data and analytics strategy"' and that doing so requires a strong, top-down approach steered by a leaders who are able to get the commitment of the rest of the firm. Therefore it can be argued that the people, organisational, technological and process challenges they identified as impacting the ability of an organisation to create value from big data and analytics cannot be overcome without first overcoming any leadership challenges which may exist.

Kiron and Shockley (2011) noted that organisational factors and cultural commitment are important predictors of whether an organisation will be able to leverage value from analytics. More articulately, Lavalle, Lesser, Shockley, Hopkins and Kruschwitz (2011) cited executive support, culture and governance as some of the obstacles to overcome to be able to extract value from big data and analytics, once again identifying top-down leadership challenges.

According to Rosseau and McCarthy (2007), the successful implementation of evidence-based management, in this case led by data, lay in the hands of leadership that embrace the practice to be able to make it work. They noted that to date, managers have been slow in augmenting intuitive decision making with scientific evidence as a routine part of decision making (Rosseau & McCarthy, 2007). They found that the barriers to leaders adopting evidence-based management as daily practice included culture and business history, being held accountable to provide evidence in support of the decisions made, and not having the time and resources to evaluate its successful application created.

The research undertaken by McAfee and Brynjolfsson (2012) identified leadership as one of five challenges to becoming data- and analytics-enabled. They note that leaders faced challenges in developing new decision making patterns to enable them to lead decision making with data and to sub-serve their own intuition to data. They argued that by visibly adapting data-led decision making, leaders set the tone for the acceptance of data-led decision making that will filter through the organisation. At the same time, leaders should retain their vision and remain attuned to their stakeholders, including customers and shareholders, to enable them to identify and pursue new opportunities and offerings highlighted from the use of data and analytics. Execution could be challenging in practice, where the processing of data, without the right vision, could lead to insights that never translate into improved products, services or products and therefore never create value.

Willmott and Dewhurst (2014) stated that leaders need to lead the way: A behavioural shift is required by leaders in order for their organisations to have any hope of reaping the competitive benefits of early adopting big data and analytics through improving the speed and quality of strategic decision making. Importantly, leaders will need to apply their ability to handle ambiguity. Unlike situations where outcomes are known, the big data and analytics world requires leaders to foray into the unknown as they synthesise outputs from disparate, emerging data sets to evolve proof-of-concepts into business practice. They need to resist the temptation to engineer an outcome that the data does not support.

Bell (2015) noted that for data and analytics to become a strategic asset, leadership support is critical. Willmott and Dewhurst (2014) found that leaders who are at the forefront of the big data and analytics across various industries had to adapt their managerial style to integrate it. To do this, they believe management needed to unlearn many years of organisational development, to "let go" and transform their role. They could do this through the probing questions they pose and the manner in which they tackle output by algorithms. At the same time, management needed and focus on the human elements of management such as inspiring rejuvenation of an organisation. This dovetails with the arguments put forth by Rousseau and McCarthy (2007) around the revision of management education to be more relevant to the modern, data rich economy. Leaders are realising that having big data and data scientists on board to process said data is not enough to create value from it. Somehow, the insights gleaned need to come to fruition as improved products, services and processes. Willmott and Dewhurst (2014) argued that the one-way flow of information to the top, to be controlled by leadership, has come to an end due to the power in crowds. This change in the dynamic of information flow and access poses an uncomfortable change to leaders.

Brown, Court and Wilmot (2013) pointed out that due to the constraints leaders already face, they are hard pressed to respond and develop the requisite capabilities to undertake the organisation-wide change required to extract value from big data and analytics. Therefore, even though some of the capabilities discussed thus far may not be new leadership skills and concepts, the context and speed at which these skills need to be acquired and exercised add a layer of complexity managers need to be aware of. The top-down approach of setting the right tone required to leverage big data and analytics has not been easy to implement, even though the literature suggests that a change in leadership capability and a strong leadership orientation toward actually using big data and analytics to make decisions will lead to buy-in throughout the organisation and enable the firm to use big data more judiciously in creating value. The fact that leaders are also simultaneously called to remain dynamic and build a vision and have a strategy in place to govern how big data and analytics will be part of the overall firm strategy adds to this challenge. It is evident that the harnessing of big data and analytics is a big ask of managers, who may not currently be well equipped to undertake the challenge of becoming data-led.

Key thoughts: Organisations need leadership support and collaboration to create value from big data and analytics

2.7 Data sharing, privacy, and ethics

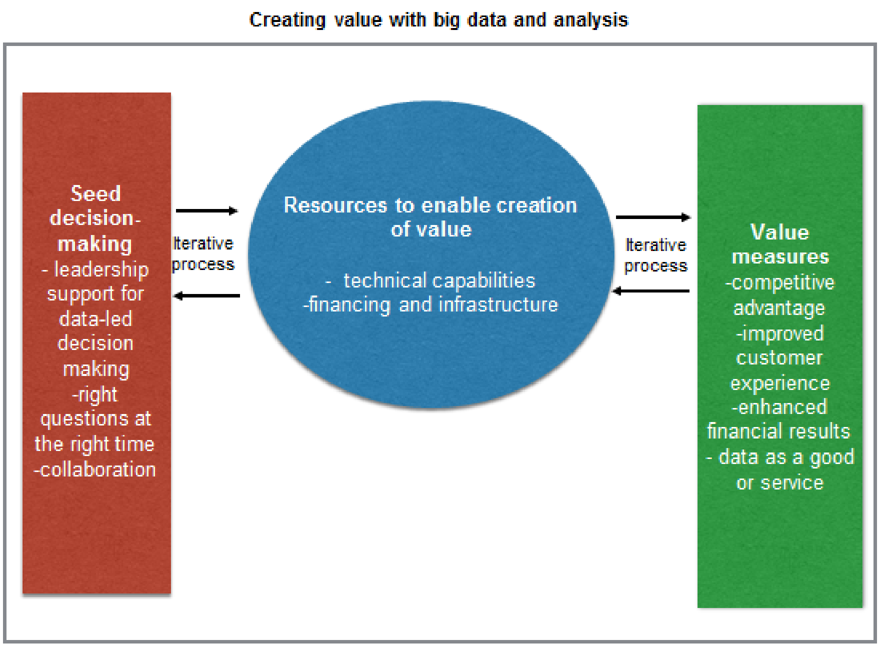
Big data and related analytics give rise to social issues. Arguably, this becomes especially pertinent when big data and analytics are packaged and sold as a product, rather than just used internally to drive decisions and develop products and services. In these cases data protection, sanitisation and anonymisation was pivotal. According to George et al. (2014), approaches needed to be formulated by explicitly considering the issue of confidentiality and anonymity of data being breached through triangulating multiple bodies of data. Data security to limit unauthorised access, the management of privacy and other rights, and ensuring that data is not misused has become pivotal. Boyd and Crawford (2011) echoed the view that the publiRP5ess of data does not equate to permission for unconstrained use of that information, and mentioned that the use of social information can cause ethical issues. Seddon and Currie (2010) noted that regulation does not always keep up with advancing technology. This creates the opportunities within which intentional or inadvertent data breaches can take place.

Although challenges in data sharing, privacy and ethical considerations are acknowledged as serious issues that organisations looking to leverage big data and analytics will face, it is in and of itself a topic for future studies and is therefore excluded from the scope of the current study.

2.8 Conceptual Model

Based on the key thoughts presented at various junctures in the literature review, the researcher has compiled the diagram, presented as Figure2.8.1 to illustrate what needs to be in place for an organisation to create value from big data and analytics.

Figure 2.8.1 Conceptual model



As illustrated in Figure 2.8.1, the concepts highlighted from the literature review have been organised into a process to note the building blocks that to be in place in an organisation wishing to embark on a big data and analytics journey: the right support for data led decision making, the enablers of data and analytics and an end goal: value. These concepts have informed the research propositions set out in Chapter 3.

2.9 Conclusion

The above literature has illustrated the evolving nature of the concept of big data and analytics, as well as the topical nature of the pursuit of value through the use of big data and analytics. Big data and analytics have been acknowledged as being important to all parts of all organisations, and big data principles are being adopted across many industries as organisations strive to harness this new approach to create an information edge that could position them ahead of their peers. It is however clear that organisations have not mastered the art of creating value from big data and analytics, with numerous challenges identified from the perspective of leaders. This has suggested that an opportunity exists to explore this topic further, both from an academic and more practical business perspective. By establishing key elements of the process through which organisations create value through big data and analytics, a reference point is created for organisation embarking on the journey to be able to do so with a greater awareness of how to prepare itself and what to expect. These key elements have been captured in a conceptual framework (Figure 2.8.1). The specific research propositions that have been identified as worth exploring, based on this framework are presented in Chapter 3. These research propositions are based on the process required by organisations as it implements the wide range of changes required to overcome the new challenges presented by big data and analytics.

Chapter 3: Research propositions

3.1 Introduction

The purpose of this qualitative research was to understand how to leverage big data and analytics as a source of value, as seen from a leader's perspective. The literature led to the formulation of the research propositions, which were used to explore the central phenomenon or topic (Creswell, 2012).

Figure 2.8.1 sets out the components of the process for creating value from big data and analytics, as informed by the literature presented in Chapter 2. It provides the basis for the propositions that follow.

3.2 Research proposition 1: Decision making

To create value from big data and analytics:

- Organisations need **leadership** that **support** the use of big data and analytics in decision making;
- Organisations need to have the ability to **ask the right questions** at the right time;
- Organisations need to have **collaboration** between their data and analytics functions and the rest of the business.

3.3 Research proposition 2: Resources

To achieve value out of big data and analytics, organisations need to have the enabling **resources**, which include:

- **technical** skills/capabilities;
- **financial resources** and **infrastructure.**

3.4 Research proposition 3: Value measures

To create value with big data and analytics, organisations need to be able to **articulate** what **value** it is looking for, as this enables the tracking of success through **metrics**.

Chapter 4: Research methodology

4.1 Introduction

As illustrated by the literature reviewed and presented in Chapter 2, big data and analytics is a relatively new concept and although research in its field exists to date, an extensive body of research into the leadership perspective on creating value from big data and analytics was not found. As a result, to gain a deeper understanding in exploring this perspective, the study was undertaken using an exploratory, qualitative approach. A qualitative or interpretive approach placed emphasis on the participants of the research, and focussed on gaining an understanding of the unique perspective they offered. The exploratory nature was aimed at discovering general information on the topic of creating value through big data and analytics in addition to that provided by existing studies (Saunders & Lewis, 2012). As is typical of an exploratory study, initially there was no firmly proposed answer; simply research propositions which the researcher undertook to address (Saunders & Lewis, 2012). Even though no rigid theory was espoused, existing literature was used to formulate these propositions and the approach of this study was aimed at confirming or modifying these propositions and hence, the research was deductive in nature (Saunders & Lewis, 2012). Data was collected through interviews and assessed to determine if it supported the research propositions posed. In doing a qualitative study, knowledge was gained from the personal views expressed by participants and the meaning they ascribe to their observations (Creswell & Miller, 1997).

4.2 Data collection and measurement instrument

As a qualitative, exploratory approach was decided on, the appropriate data collection method needed to be selected. To gain an understanding of leaders' experiences, direct knowledge needed to be obtained from leaders and therefore the researcher opted for detailed, semi-structured interviews applying the long interview method (McCracken, 1988) to allow for a comprehensive approach to primary data collection. The interview method allowed participants in the research to tell their stories (Creswell & Miller, 1997). Thus, the interviews offered flexibility and breadth as an information discovery and collection tool, which made it more useful in this research than alternative methods.

Saunders and Lewis (2012) noted that a semi-structured interview was where the researcher had a list of questions to ask, but also had flexibility in how the questions were raised. The semi-structured nature of the interview proved helpful as it made allowances for the fact that:

- The researcher was often unsure of the nature and scope of responses the participant would provide, and questions may have been deemed irrelevant to the participant, and therefore may have been omitted.

- The order in which questions were raised could be varied based on the participant's responses.
- The questions were complicated.

The researcher found that all participants approached and handled the interview differently. Some were very structured in providing responses, while others spoke more generally and appeared at times to deviate somewhat from the topic at hand. The researcher believes that the semi-structured approach, through its more conversational rather than interrogative style facilitated the free sharing of information better than alternative methods, by placing more control over the interview in the hands of the participant. It also provided the researcher with flexibility to dig as deeply as needed to gather information in support of the research objectives. The researcher was also able to explore areas mentioned briefly by the participants and hence the participants were also afforded time to reflect on their thoughts (Seddon & Currie, 2017).

The long interview approach enabled a thorough exploration of the mind of the participant (McCracken, 1988). The long interview methodology enabled the researcher to use their own culture to formulate better questions, to be more cognisant in the interview, thus adding to the richness of the study (McCracken, 1988). In this way, more data was collected, while it still remained manageable to work with.

Interviews were conducted with individuals, rather than groups of individuals, as it permitted the researcher to form a relationship with the participant and gain better insight into the participant's perception of creating value with big data and analytics. An individual approach resulted in the participant not having to consider filtering their responses due to fear of how it would be perceived by other participants. This also enabled the researcher to build rapport with the participant more easily and foster greater trust, and this further improved the extent to which participants could speak freely and truthfully, compared to a scenario where other individuals are present. Face-to-face interviews also permitted non-verbal communication, such as body language, to be noted and the researcher could adapt the line of questioning to be responsive to this. The interviews were also be conducted in the workplace or an environment in which the participant was comfortable, as this put them at ease and proved conducive to greater sharing of insights (Creswell & Miller, 1997). At the same time, this method did not result in an undue intrusion into privacy of the participant.

The interviews were conducted on an anonymous and confidential basis, meaning that the name of the participant and their organisation was not be linked to specific responses in the research report, nor shared with anyone who is not involved in the academic research (Creswell, 2012). The researcher asked the participant to provide written consent both to be interviewed, and for the interview to be audio recorded, and clarified the confidentiality and right to withdraw from the interview at any stage. An example of the consent form is included in Appendix 1.

The researcher used a pre-prepared list of questions to guide and provide structure to the discussion. However, the nature of the research methodology allowed the participants to guide discussions, and the questioning process was dynamic, with no two interviews having the same question set. The researcher did not lead discussions, nor try to guide it to formulate consensus of opinion, thus providing more credibility to the study (Creswell, 2012).

The first interview was used as a pilot to ensure the questions were open ended and would lead the researcher to the necessary data (Saunders & Lewis, 2012). During the process of data collection, the questions were also updated to delve in more detail and accommodate the evolving nature of the study. The researcher became more comfortable with the process as the interviews progressed and was able to take greater steer from subsequent discussions for adding questions dynamically during the interviews. The result was that by interview four, the researcher no longer stuck to the questions verbatim and tried simply to ensure the topics were explored during the discussion. If a topic was raised spontaneously by a participant, the researcher simply let the discussion flow in that direction, as opposed to “parking” that topic for later discussion. The initial and updated discussion guides are included in Appendix 2.

Interviews were recorded using a digital voice recording device. The researcher did not take detailed handwritten notes, and instead engaged in eye contact and was therefore able to observe and take cues from the participant’s body language. The researcher did take notes of points or areas for further questioning at a later opportune moment. The researcher created back-up copies of the audio recorder on an external hard drive as well as on a cloud-based application. Due to time constraints the researcher enlisted a reputable, professional transcription service to assist with converting the interviews from audio to text data.

4.3 Population

Saunders and Lewis (2012) described a population as “the complete set of group members” (p. 132). The population included all organisations operating in South Africa that were using big data and analytics. Additionally, the organisations also needed to have senior decision makers or leaders of the big data and analytics programmes (such as chief information officers, chief digital strategists, chief innovation officers).

4.3.1 South African context

Plessis and Boon (2004) note that the South African business environment is different from other countries and that this needed to be catered for in the implementation of knowledge management systems. The same can be argued for this specific study into big data and analytics, which is somewhat related to knowledge management systems. Searches performed by the researcher revealed an apparent shortage of academic research in highly rated publications on big data and analytics within the South African context.

Baller, Dutta and Lanvin (2016) found that South Africa's digital transformation was driven by commercial usage of information, however the innovation environment was deteriorating; potentially at the hand of expensive mobile and broadband services. Ndlela and du Toit (2001) found that South African organisations recognise that knowledge management could provide a competitive advantage, yet a lack of leadership and a shortage of time and resources posed challenges to knowledge management systems. This paved the way for a challenging landscape in which leaders gear themselves up to extract value from big data and analytics. These contextual nuances supported a case for the specific study of a leader's perspective of big data and analytics in South Africa.

Corte-Real et al. (2016) conducted their study in the European context and cited a need for big data and analytics studies in other countries and industries. They posited that understanding gained from different external environments will advance the body of knowledge on using big data and analytics. This supported the exploration of this topic within the South African context.

4.3.2 Financial services

The financial services sector is constantly changing and technology is a key source of competition (Seddon & Currie, 2017). Financial service organisations have rich sources of data which, it has collected in order to manage the risks it exposes itself to, as well as to comply with regulatory requirements for example, anti-money laundering provisions. George et al. (2016) noted that banks are using that data to apply behavior modelling and leverage data science in the risk and compliance functions on a real time basis.

DalleMulle and Davenport (2017) further noted that banks operate in dynamic and often highly competitive industry and do not only use their data tactically, but that they also often used their data strategically. Consequently, the banking sector is set to benefit very strongly from big data as long as barriers to its use can be overcome.

Based on the natural inclination for the financial services toward being data rich, and the anticipated maturity in leveraging big data to create value, this research was explored within the context of this sector.

4.4 Sampling method and size

The participants of the research were identified through purposive, non-probability sampling as a complete list of companies applying big data and analytics within the South African context did not exist (Creswell, 2012). For this reason, a combination convenience sampling was applied through leveraging the researcher's existing networks to identify the participants. Each participant in the research interviews formed a sampling unit.

According to Mason (2010), a variety of factors could influence the sample size in a qualitative study, including data saturation, and consequently many researchers were reluctant to specify guideline sample sizes. McCracken (1988) noted that the purpose of the qualitative process is to do an intensive study and therefore noted that a smaller sampler size is preferred for a more in depth study, and proposes a sample of eight interviews. However, the researcher aimed to conduct 15 interviews, in line with the minimum sample applied in 80% of PHD studies (Mason, 2010). Mason (2010) noted that data saturation is influenced by a variety of factors, such as the heterogeneity of the population, resource availability (which includes time), the expertise of the participants and the scope and nature of the study. These factors made it difficult to predict the precise sample size upfront. Even through the researcher aimed for 15 interviews, by the ninth interview no substantially new points were being raised by participants, and given the range of interviews conducted to that point, no further interviews were pursued. Part of a reason that saturation occurred at this point is that the sample was drawn from the financial services industry, which led to an element of homogeneity in the sample.

The study was initially aimed at South African banks; however, challenges arose in securing interviews with participants who were willing to be interviewed once they found out that the researcher is employed within the office of a member of executive management of one of the four large South African banks. As a result, the sampled participants were employed in organisations across the financial services industry, including the insurance subsector. One participant was employed in consulting organisation, and his role was focussed on consulting in the financial services industry. The majority of participants were employed in banks, and this biased the sample toward the banking industry. Table 4.4.1 outlines the interview dates, identifiers of participant, their position or designation within their organisation and the industry they operate in.

Table: 4.4.1 details of research participants (RP)

Date	Participant	Position	Industry
24 July 2017	RP1	Chief information officer: group	Banking
25 July 2017	RP2	Head: pricing & analytics - Forex	Banking
1 August 2017	RP3	Head of analytics & specialist pricing	Banking
10 August 2017	RP4	Director: risk advisory: data and analytics (FS sector head)	Consulting
10 August 2017	RP5	Chief executive officer (CEO) (divisional)	Banking
15 August 2017	RP6	Head: pricing & analytics	Banking
16 August 2017	RP7	Chief executive officer (CEO)	Insurance
17 August 2017	RP8	Chief information officer	Banking
6 September 2017	RP9	Chief information officer	Insurance

4.5 Data analysis

According to Oliver, Serovich and Mason (2005), transcribing is a powerful act of representation. The transcription was performed by a professional service enlisted by the researcher. The researcher reviewed the transcripts against the audio recordings to ensure that the transcriptions were done accurately and completely. A denaturalist approach to transcribing the interviews was applied. Under this approach the focus was on the substance of the interview – emphasis was not placed on a verbatim account of the interview, but the interview was still transcribed to provide a faithful rendering of the interview (Oliver et al., 2005). This approach ensured that the researcher captured the essence of the interviews by omitting interview noise, such as pauses, works like “uhm” and repetition of words as people formed their thinking, or errors in speech, and wrote the interview in a more flowing manner to convey meaning in the reading that would otherwise have gotten lost without the voice over. As recommended by Oliver et al. (2005), the researcher also reflected on the transcribed information to ensure that the detail captured honoured both the research process and the participant’s voice.

Each interview was transcribed in a consistent manner and this enabled the subsequent coding process of organising the content and highlighting meaning to be to be undertaken with greater ease. The researcher ensured that all information which could be used to identify the participants or the organisations they work at were removed from the transcriptions. No information such as age, gender or ethnicity that was not relevant to the study was asked or recorded. The researcher analysed the data gathered from interviews personally to gain insight on challenges faced in using big data and analytics to create value.

A qualitative data analysis software program, ATLAS.ti, was used in the coding and categorisation process. Sinkovics, Penz & Ghuari (2008) posit that software programs aid in substantiating the analysis and interpretation of interview data. This therefore enabled the researcher to undertake a more efficient and systematic analysis of the interview data.

4.6 Method of analysis

Coding is an ongoing process of interpretation and examination of data from different perspectives, and is considered to be probably the most crucial step in the analysis (Sinkovics et al., 2008). Each interview was transcribed and coded soon after the interview to allow for the timely identification of data saturation.

A combination of the inductive and deductive approaches was followed. The approach allowed a more robust interpretation of the information gathered, and thus provided the most appropriate way to understand the perceptions of challenges in using big data and analytics to create value.

The researcher started with a deductive approach as the researcher started with a close-ended view of themes which were used to collect and analyse the data (Creswell, 2012). Based on the literature review, the researcher identified themes of creating value through leveraging big data and analytics prior to commencement of the coding process. These themes were used as a guide through which to analyse the responses provided during the interviews. These themes were identified in accordance with the steps proposed by Saunders and Lewis (2012), whereby:

1) Meaningful codes were developed.
2) The right pieces of text data to which to attach the codes were identified. The right piece varied and may be a line of text, a sentence or a paragraph of a response.
3) The codes formulated in 1) above were attached to the pieces of data identified in 2) above.

The codes that were identified in 1) were identified from the literature review and the initial coding table used is presented in Appendix 3.

Given the exploratory nature of the study, the researcher catered for the discovery of additional codes by following an inductive approach as well. In this way any other, new codes and themes that surfaced during the course of the interview analysis, potentially specific to the South African environment, were also named and recorded. This approach highlighted the additional challenges that were identified by the participants so that a narrowing approach could be used to identify additional themes through appropriate aggregation (Creswell, 2012; Saunders & Lewis, 2012). The researcher also expected that the codes identified upfront and during the study would change to remain meaningful as the research progressed (McCracken, 1988; Saunders & Lewis, 2012).

The research followed a mono-method approach as only qualitative data collection was employed (Saunders & Lewis, 2012). Furthermore, a cross-sectional research design was followed, as interviews were conducted in one period only (Saunders & Lewis, 2012).

4.7 Research limitations

Many critics challenge the reliability of qualitative studies (Shenton, 2004). Reliability is the extent to which the data collection and analysis process would produce consistent results (Saunders & Lewis, 2012). Validity is also a challenge, and refers to the extent to which the data collection method, in this case interviews, can accurately measure what it is intended to; and the extent to which research findings are a true reflection of the area of research. The high level of subjectivity that goes into the qualitative data collection and analysis process causes many of the concerns around validity and reliability. This subjectivity is due to the fact that the researcher is in essence n inextricably part of the data collection process as they function as the data collection instrument and therefore are cannot fulfil their role without allowing their own experience and intellect to inform the research (McCracken, 1988). Even though this may offer richness to the study as it can permit a transfer of experience from the researcher to the words of the participant to create a richer understanding (McCracken, 1988), the researcher exercised rigour in undertaking the research and ensured that the study met the criteria of transferability, dependability, conformability and credibility, and that other related influences on the study were identified, and managed where possible. Part of this process was to ensure that the researcher used a consistent approach and procedures in the research, as making alterations to the procedures followed could introduce bias into the study (Creswell, 2012).

Researcher's influence on the research

Qualitative studies may inherently be subjective and to cater for this, researchers need to reflect on and name their own preconceptions, values, and assumptions (Creswell, 2012). In this case, the researcher therefore must acknowledge that they are employed in the financial services sector, within the office of an executive of one of the four large South African banks and had limited insights into some of the big data and analytic dynamics of the financial services industry, which may have caused some emphasis to be placed on certain types of questions, and may have resulted in a breach of non-directive questioning (McCracken, 1988). However, the researcher's professional background was in accounting and the researcher was new to the financial services role (employed May 2017). This offered some mitigation of this impact. Additionally, due to the personal involvement of the researcher, the researcher's history, experience, and values also influence the manner in which the research is analysed and the subsequent findings and conclusions (McCracken, 1988; Saunders & Lewis, 2012).

Transferability of the research

To ensure that the research findings are used within the right context, and not inappropriately generalised, the research report includes detailed contextual information about influences introduced by the sample organisation and individuals, as per the previous sections. This enables readers to gauge the extent of transferability of the research outcomes (Shenton, 2004). The researcher remained cognisant of the fact that providing too much background information on entities may compromise confidentiality. Given the range of companies in South Africa, the researcher believes this has been achieved. Information on the context of South Africa has also been provided to cater to international attempts at transferability.

Dependability of the research

According to Shenton (2004), the dependability of a study is derived from the ability of another researcher to undertake the same study, even if not to arrive at the same research findings. The researcher ensured this by providing detail on research design, and clear, specific detail of what was done in the field to enable other researchers to follow a similar approach.

Confirmability

Confirmability is the extent to which research findings reflects the experiences and views of the participants, rather than the capturing the interpretive predispositions of the researcher (Shenton, 2004). Sinkovics et al*.* (2008), added that to enhance validity, information must be systematically and consistently collected. To improve the validity of the study, vivid detail was provided to present contextual information by using rich descriptions of the interviews, including the provision of snippets of the interaction throughout the documentation of the research findings (Creswell & Miller, 2000). The reported findings therefore went beyond a sterile description of facts and also outline detail to make the reader feel as if they experienced aspects of the research. This robust description will also enable the reader to better assess the transferability of the study (Creswell & Miller, 2000).

Credibility of the research

Credibility is about the researcher providing a reflection of the realities of participants (Sinkovics et al., 2008). The personal means in which this study was conducted resulted in the researcher needing to acknowledge their inseparableness and the resultant potential influence on the research process, and the impact this may have on the study. Qualitative studies need to aim to prove the credibility of their studies for it to be relied on by consumers thereof (Creswell & Miller, 2002). Personal views and interpretations held by the researcher were clearly expressed as such (Creswell & Miller, 1997).

Use of the long interview method also positioned the researcher to be aware of her own views and beliefs, and to understand and account for these subjective factors during the study (McCracken, 1988).

Other limitations of the study

The research has the following limitations:

- Time was a constraint of the study (McCracken, 1988). Access to executives in blue-chip data driven organisations was limited due to the short timeframe for data collection inherent in the study.
- The sample selection methodology of convenience sampling resulted in a sample of similarly predisposed individuals, and hence limited the variety of insights and richness of the research findings.
- The research process is qualitative, and findings were influenced by the personal interpretations and specific facts and circumstances experienced by the participants. The findings are generally not applicable to other situations (Shenton, 2004).

4.8 Ethical considerations of the study

Various ethical consideration and practices were undertaken during the research. This enabled an interview environment of trust and a more informative dialogue. All of the participants were comfortable to be interviewed in English and therefore a translator was not required.

At the outset of each interview, the researcher provided the participant with a printed consent form and explained its content. The researcher then allowed participants time to read and then sign the document. In addition to not asking information such as age and race which are not relevant to the study, the following practices were also observed (Creswell, 2012):

- The participants were briefed on the purpose of the study and the fact that they may cease their participation therein at any point in time.
- The researcher was transparent and answered all questions posed by the participant honestly; bearing in mind the need to maintain confidentiality in cases where the participants enquired as to other interviews the researcher had conducted.
- The researcher did not disrupt the research site or act disrespectfully or inappropriately when visiting the offices of the participant.
- The researcher maintained the confidentiality of all information gathered, as outlined in this report, mainly through sanitising the transcribed interviews used in the research and referring to participants only by a pseudonym in the research. In addition, one request to be "off the record" was noted during and interview and this content, although recorded, was omitted from the transcribed interviews and analysis (Creswell, 2012).

Chapter 5: Research results

5.1 Introduction

The researcher conducted nine interviews that collectively offered numerous, varied insights into practical experiences in creating value from big data and analytics within the financial services sector.

This chapter commences with a summary of the interviews undertaken, providing details of the individuals interviewed so as to provide context to the results. The chapter also contains an account of the procedures performed by the researcher in assuring the accuracy and validity of the interview, and outlines the transcription and analysis process. A combination of content and thematic analysis was performed. The chapter contains the interview analysis within the context of the research propositions put forth in Chapter 3. In addition, themes which were discovered inductively are also discussed.

5.2 Data gathering and analysis

5.2.1 Summary of the interviews conducted and the interview method

Initially, the researcher wished to conduct up to fifteen interviews. However, after undertaking the first eight interviews, the researcher did not gather further significant, new information. One further interview as done as it had already been committed too, and, as a result, nine interviews were conducted at which point data saturation had occurred and this then became the guide to cut off the interviewing process (Mason, 2010). This was also in line with the guidance espoused by McCracken, who noted that in most cases a sample of eight interviews would be adequate as a qualitative approach focussed on the intensity, rather than the extent of research (McCracken, 1988). Although there were more interview opportunities available, it was decided that they would not be pursued, as the ability of the researcher to analyse and provide a rich, detailed analysis diminishes with each interview added (Creswell, 2012). Hence, to preserve the complexity and avoid giving a superficial account of information gathered, only nine interviews were conducted.

Table 5.2.1.1 contains details of the research participants, illustrating their position as leaders within their respective organisations and their involvement in the financial service sector within South Africa as outlined in Chapter 4. Three participants identified themselves as chief information officers (CIOs), two as chief executive officers (CEOs) and three as divisional heads within their banks, reporting directly into c-suite executives. A director in a large consulting practice focussing on the financial services sector was also interviewed. The inclusion of different leadership roles and levels allowed the researcher to explore a greater array of leaders' perspectives. The researcher secured interviews from her own social connections, rather than by way of the networks of the participants.

Per the below table, the interviews resulted in a total of 448 minutes of audio, or just under seven and a half hours of interview recordings, averaging 49 minutes per interview. The content of these interviews resulted in transcripts of a total of 66 884 words, or an average of 7431 words per interview. The length of interviews allowed participants sufficient time to delve into as much detail as they desired on the topics discussed, while still being mindful of their time constraints as very senior leaders within their organisations. In addition, five of the participants were in the banking sector, giving this sub-segment of the financial services industry the loudest share of voice. Based on their job title, the majority of the participants (six out of the nine) were "chief…" or "director" and operated at an executive level in their organisations and thus offered the strategic perspective a greater amount of airtime. Three participants held the job title "head of..." and hence are occupied at a management level in their organisations, offering a more operational perspective.

Table: 5.2.1.1: Details of interviews and research participants

Number	Date	Participant	Position	Industry	Length (min)	Word count
1	24 July 2017	RP1	Chief information officer: group	Banking	54.25	7956
2	25 July 2017	RP2	Head: pricing & analytics	Banking	49.37	5737
3	1 August 2017	RP3	Head of analytics & specialist pricing	Banking	39.20	6142
4	10 August 2017	RP4	Director: Risk advisory: data and analytics (FS sector head)	Consulting	49.19	7465
5	10 August 2017	RP5	Chief executive officer (divisional)	Banking	58.02	7848
6	15 August 2017	RP6	Head: pricing & analytics	Banking	48.41	8319
7	16 August 2017	RP7	Chief executive officer (divisional)	Insurance	52.24	7542
8	17 August 2017	RP8	Chief information officer	Banking	47.67	6940
9	6 September 2017	RP9	Chief information officer	Insurance	49.20	8935

Number	Date	Participant	Position	Industry	Length (min)	Word count
	Total				447.55	66 884

The researcher obtained ethical clearance on 22 June 2017 (see Appendix 4) and started setting up interviews thereafter. Interviews proved difficult to schedule due to the availability of the senior leadership targeted and as a result were undertaken over a six week period starting on 24 July 2017. The interviews were framed by the discussion guide which was developed ahead of all the interviews. All interviews were conducted in person, in a location proposed by the participant – mainly in a meeting room on their company premises. In one instance the interview was conducted in a cafe on the premises of the participant, and a significant amount of background noise occurred during this interview.

All interviews were recorded digitally on a voice recorder and were downloaded onto the researcher's laptop and thereafter onto a cloud based storage application. This was done within hours of conducting the relevant interview. The interviewer did not take extensive notes during the interview, and preferred to maintain eye contact and pay attention to the body language of the participant as it is key, in interviews, to listen with great care to pick up nuances such as avoidance or discomfort (McCracken, 1988). The notes that were taken related to interesting points the researcher wanted to explore further at a later stage in the interview, if the participant did not expand on it naturally.

In advance of the first interview, the researcher undertook research on the best practice methods to prepare for and conduct an interview (Creswell, 2012). After the interview was conducted the researcher reflected on the interview technique applied to assess how well the technique had been followed, especially as it relates to open-ended questioning. In addition, the appropriateness of the discussion guide was considered and necessary revisions were made, although these were not major.

The second interview followed shortly after the first. The second interview respondent noted that the research questions were very broad. As a result, after the second interview the discussion guide was updated to enhance the flow of discussion and to make the questions more specific to the research propositions, taking cognisance of the requirement outlined per McCracken (1988) to keep the interview broad and allow discovery of the participant's unique story instead of guiding the discussion to obtain preconceived responses. In this way, detailed but diverse responses were obtained. The researcher decided to imbed some wording around the general or broad nature of the questions in the introduction of interviews to manage perceptions around the type of questions which were to follow. Although the interviews all followed different sequencing and wording of questions and added different case-specific questions (some of which were raised in more than one interview), the base discussion guide was not revised again. The added preparatory narrative and the altered discussion guides were used in subsequent interviews and the researcher was comfortable that the introduction and discussion guide was adequate. The initial and revised discussion guides are included in Appendix 2.

During the currency of this study, the researcher was employed by a bank which forms part of a larger financial services group. Only one of the participants was known to the interviewer before the interview was undertaken. However, four of the participants were employed by the group of financial service organisations in which the researcher is employed. This resulted in these participants using language which implied familiarity with the researcher such “us” and “we”. In addition, references to an assumed amount of knowledge on some activities within their organisations were made during the interviews. The group is highly federated and various organisations and segments within the group operate as separate brands with discrete strategies, boards, and executives directors. A strong owner-managed culture exists and as a result each segment is effectively operated as a separate business. Based on this operating model, the organisations were included within the sample as they offered different perspectives.

All interviews were conducted face-to-face with only the participant present and after building initial rapport they appeared comfortable to share their perspectives openly. All interviews were undertaken in private settings, barring interview three, which was conducted in a cafe at the premises of the participant. The researcher found that this participant was more reticent than the previous two participants and the responses to questions occasionally seemed incomplete and sometimes lacking in detail. The researcher had to do more probing than in the prior instances. This could have been due to the participant’s personality, but also potentially due to the venue utilised as the participant might have felt he could not speak as freely as he would have in a more private setting. The cafe also became disruptive at some point, which somewhat disturbed the flow of the conversation. Two interviews were interrupted- one through an intrusion by and external party into the meeting room, which disrupted the train of conversation, while the other interview was disrupted due to a telephonic call that interfered with the recording. This resulted in approximately five minutes of the interview content was irretrievably lost.

All participants were handed a consent form at the onset of the interview, and the researcher explained its content. One participant expressed surprise at being recorded and the researcher explained the purpose of the recording as well as how the recording would be managed, and then confirmed that the participant was comfortable with this. The participant then signed the consent form. All consent forms were saved electronically as part of the body of evidence to support this study. Overall, participants were comfortable to sign the consent form and to be part of the study.

At the start of all interviews, the researcher ensured participants were aware of her status as an employee in the office of an executive member of a large bank. This may have caused reservation in some responses as participants remained protective of their intellectual property and cognisant of not revealing too much future strategic information. However, this did not deter their general willingness to discuss their big data and analytics journey.

5.2.2 Interview transcription and verification.

In the interest of time, a transcription service was used. The researcher had a detailed briefing session with the transcription services to discuss the denaturalist transcription methodology to be used, as outlined in Chapter 4, and to agree on the format and timing of required feedback. All nine interviews were transcribed by the same transcription service provider. Once transcribed interviews were received, the researcher followed a systematic approach of listening to the audio file of each interview and comparing it to the transcribed version thereof to ensure that the two were consistent (Spiggle, 1994). This process took around two hours per interview. As part of this process, the researcher corrected any transcribing errors which may have occurred, such as where “Hadoop” was captured as “adoop”. The researcher was able to update the transcript based on a better understanding of the topic of conversation, context and memory of the conversation. In a few isolated cases, the meaning was not immediately evident and the researcher had to rewind and repeat the relevant section of the conversation to decipher meaning. Speech that remained inaudible was noted as such. The researcher also aimed to provide better structure to the flow of transcribed interviews by adding punctuation and breaking sentences into shorter parts where this was not appropriately done by the transcriber and led to confusing text. The researcher did not focus on removing filler phrases such as “you know”, “so”, or “ok”, or colloquialisms such as “cool”. This was not removed because while it does not add substantial information, it also does not detract from the interview and is a closer reflection of the character and style of the participant and the reality of how they expressed themselves. The transcripts were sanitised to remove references to the names of the organisations at which the participants worked. Names of other companies mentioned were not removed, unless it was a reference to a competitor. For example, if the participant mentioned a company in the context of a banking client relationship, the name was removed. A generic reference to, for example, a retailer, was not removed from the transcripts. Industry-specific abbreviations used by participants were updated to be written in full – for example, "FX" was changed to "foreign exchange". The researcher also deleted all information relating to an instance where the participant made it known that they wished to be "off the record" but where the audio recording was not stopped for practical reasons of not disturbing the flow of conversation, so that it may not be included in the analysis. Interviews were reviewed in the sequence in which it was undertaken over a nine day period. Once the comparison to the audio file to the transcript was completed, the researcher re-read the interview to ensure no spelling or confusing phrases or sentences remained. This process provided an initial view of similarities and differences between the interviews. However, no formal analysis was undertaken during this preparation process.

5.2.3 Coding and analysis of transcripts

The analysis of the interviews was undertaken in ATLAS.ti, a qualitative data analysis software program. The researcher undertook the analysis of all the transcribed interviews personally.

Transcript coding

The coding process generally followed the process outlined by Creswell (2012) for analysis using qualitative computer programs. As applied in this case, the relevant steps applied, in sequence, were:

- ATLAS.ti was selected as the software program to use due to its ability to store, data, organise data, assign codes, and perform searches of the data.

- A Microsoft Word version of each transcribed interview was loaded onto ATLAS.ti. The transcripts were named according to the sequence of the interviews conducted, with the identifier of the participant added as a post script. In this manner, any quotes or codes could be linked back to the person who had made it. For example, the first interview was titled "Interview 1_RP1".

- The researcher went through each file in ATLAS.ti and identified words, sentences or paragraphs containing key ideas that the participant was voicing. These blocked pieces of text that conveyed an idea were assigned a code label. The naming of the code labels was informed by the initial coding table, as presented in Appendix 3. This process was followed until the entire document had been assigned code labels where relevant. This was done for each interview's transcript and interview transcripts were coded in the sequence in which the interviews were undertaken.

- After blocking and assigning code labels to text, codes were viewed in the "code manager" function of ATLAS.ti and the naming was reviewed to ensure it was appropriate. All text associated with a particular code was also reviewed to ensure that the code captured similar concepts and the code names were updated where required, in order to be better descriptors.
- Thereafter codes were aggregated using the "group" function within ATLAS.ti. The groupings represented a few broad themes or categories. These codes which formed part of each groups were given a specific colour to enable easier further identification and analysis. Creswell (2012) suggested that only five to seven themes be identified to support a more detailed analysis of a few concepts rather than a superficial review of a number of themes. The researcher identified 11 themes in this manner. These themes and the codes related to it are discussed in the thematic analysis contained in Chapter 5.4. Given the deductive approach of this research, the researcher identified themes and codes and captured these in an initial coding table, Appendix 3. The themes and codes were developed based on the research propositions. This coding table captures the initial codes that the researcher kept in mind as she was reading through the transcripts in ATLAS.ti. The codes that were preconceived by the researcher were not captured into ATLAS.ti in advance and instead, the relevant codes from the coding table were added manually as they were identified. The codes were not added upfront (referred to as "list coding" in ATLAS.ti), and instead "open coding" was opted for as the researcher felt that identifying and naming codes rather than selecting from a pre-set list would more easily permit additional codes to emerge. The researcher identified interesting phrases, paragraphs or sentences that strongly related to the themes identified in the literature review, as well as any quotes that immediately stood out as powerful. When the researcher needed to code a phrase, the researcher first determined which of the preconceived codes were relevant and reused any codes that had already been created within ATLAS.ti, and only proceeded to add new codes if none of these captured the essence of the phrase or sentence. Some of the codes presented in the coding table but were not used, while entirely new ones were added through the inductive process, if no existing codes proved to be useful in capturing the meaning in the selected phrase, sentence or paragraph. The process allowed the codes to develop and change as additional information was discovered during the analysis process. The researcher coded one interview transcript at a time by systematically working from the beginning to end thereof, identifying the preconceived codes as well as any new codes as they emerged in the first

pass. After all the interview transcripts were coded, the interviews were read a second time and any codes missed in the first pass were picked up. In this process, concepts that occur across all interviews were also identified. This final pass proved as a final check to determine if any exceptions or contradictions in the data emerged, in order to manage any confirmation bias that may have tainted the coding process.

5.2.4 Description of the research participants

As recommended by Creswell (2012), an attempt was made at describing the people, events and places of the research. Some of the description is based on direct information volunteered by the participants while supplemented with information that is based on inference.

Based on the literature review, demographic information was not relevant to this study. The line of questioning did not focus on exploring demographics with the participants. However, the upon reflection, the researcher came to believe that an understanding of the educational background and role or job description of participants aided in her personal understanding of the perspective they offered during the interviews and hence this information was captured in describing research participants in this section.

The researcher attempted to interview senior members within an organisation who could engage on the topic of big data an analytics. In the section below, the description of research participants are grouped based on the role they identified themselves as.

Chief executive officers (RP5, RP7)

RP5 identified himself as an engineer that holds a PhD in computer electronic engineering on statistical pattern recognition. Despite this, his current role in the bank is that of CEO of a customer serving segment – one of three major revenue-generating lines in the organisation. Although he was knowledgeable on the topic of big data and analytics, and has done public speeches on this topic as it relates to the banking sector, the interview content he provided was high level and his focus was strategic as opposed to operational in nature.

RP7 identified himself as an actuarial scientist who has spent the vast majority of his career within the insurance industry. He has vast operational knowledge and as he looks after the organisation's legacy systems, his job as CEO entails a strong technology management focus as well. This aspect was apparent in his interview in which IT infrastructure and organisational structure were key areas of discussion.

Chief information officer (RP1, RP8, RP9)

RP1 spent a significant portion of his time as CIO of a group of financial organisations. He had recently made the decision to resign the group CIO role to focus his efforts on the role as CIO of the specific organisation. The tone of his conversation was focussed on what he, at a high level, could solve across the organisations – hence he focussed on data quality. He spent a significant amount of time clarifying that different organisations have different abilities to create value from big data and that the precursor to this ability is the nature of clients served. In his view, organisations serving individuals have richer data – both structured and unstructured – and that organisations serving large corporates have less data and less latitude to apply analytics due to the structured and mandated manner of operating – there are less psychological influences to be uncovered and leveraged. He felt so strongly about this that he questioned the researcher on her understanding of the big data opportunities within corporate and investment banking via email when the researcher contacted him to set up the interview.

RP8 had not studied at a tertiary level after he matriculated but subsequently completed a master in business administration (MBA) degree. RP8 is chief information officer of a group of financial institutions and identifies himself as an information technology specialist. The discussion of big data and analytics was very general and initially this came across as a lack of understanding of the concepts. RP8 also made it clear to the researcher that he might not be the right person to talk to on this topic, but could not advise who in the organisation would be better positioned to undertake the interview. From this, and the futuristic manner in which he spoke, it seemed to the researcher that the organisation he was employed had aspirations but no major current big data and analytics projects underway – or none that he was willing to discuss. He spent a significant amount of time discussing the improvement of customer experiences and the difficulty in their business which is typically removed from the customer over the duration of their relationship and how to gain competitive advantage.

RP9 works as chief information organisation of an insurance organisation and nine years previously had worked at another insurance company which was much more advanced in its use of big data and analytics. She spoke with conviction of its use, but organisational limitations and lack of clarify of the value of big data and analytics seemed to have hampered its adoption within her current organisation.

Director RP4

RP4 is a chartered accountant and used the words “partner” and “director” interchangeably to describe his role within his consulting organisation, one of the four large global auditing and consulting practices. This therefore communicated his ownership in this business. He has spent his career consulting in technology aspects, specialising in cybersecurity and data and analytic space. For the past ten years he has focussed on the financial services industry. He represents the South African practice in international forums of the organisation and is close to global trends as well.

Heads of divisions (RP2, RP3, RP6)

RP3 identified himself as the chairperson of the data and analytics forum within his bank, reporting directly to the chief data officer (CDO), a role which was recently created within his bank as a result of the work he, his boss, and their team have done within the bank. He is an actuarial scientist by qualification. Similarly, RP2 is an actuarial scientist, and works in a division that services a group of financial services organisations. He is a team leader, and identified himself as senior management with a seat on his executive committee of his business unit. RP6 is the leader of team, and is a programmer by background. He is the only participant who did not explicitly describe himself as a member of an executive committee, although by inference, the researcher established that he is senior management within his business unit. These three individuals offered a more granular and operational level perceptive on creating value from big data and analytics as opposed to the strategic views discovered from the other research participants.

5.3 Content analysis

5.3.1 Transcript analysis through word counts

The ATLAS.ti "word cruncher", a content analysis tool that automatically performs a word frequency count, was used to draw a list of all the words that occurred in the transcripts, along with a total count of the frequency with which each word occurred. The word count was done for all the interview transcripts in total. The word count was viewed as a proxy for importance of concepts, as a higher frequency of word occurrence indicated which words were most used and therefore which concepts were most important to participants. Limitations of looking at a word count included the fact that no distinction was made between the researcher and the participant's responses, and the method therefore relied on the fact that the participant did most of the talking and that counts were mostly representative of words used by them. The other limitation was the fact that words, not phrases, were counted and hence the frequency of combined words could not be tracked. This meant that negative phrases versus individual words not be tracked; for example "poor" and "quality" are separated but have more meaning as "poor quality" in some instances. Similarly, other words that needed to be understood together were also missed, for example "decision making" would have registered separately as "decision" and "making". Irrespective of these shortcomings, the word count was indicative of words that were most used and therefore concepts that were most important in the interviews.

The full list of words across all interviews was exported into Microsoft Excel and 4297 words were listed before any further processing was performed. Thereafter, a sorting feature in word cruncher was used to arrange the words in the order of highest to lowest count. Next, the “exception list” feature in word cruncher was used to filter out commonly used words such as “the”, “a”, “you” etc. The researcher used her judgement to identify these as words which occurred frequently due to their function in language, but which did not provide substantial descriptive or contextual meaning to the interview, and were insignificant in the context of this study. The researcher removed these commonly used words until the point where the top 100 frequently used words were words of substance. This list was then exported to Microsoft Excel and sorted alphabetically to enable the researcher to group together words that should be counted as one word, such as “challenge” and “challenges” or “want” and “wanted”. These words were not limited to singular and plural forms, but also different word forms such “analytics” and “analyse”. The top 10 word group count that resulted from this process are included in Table 5.3.1.1. Given the focus of this research on the process of creating value from big data and analytics, it is not surprising that words such as data, decision, people and analytics occurred with greater frequency than other words.

Table 5.3.1.1: Word count from interview transcripts.

Word group	Count
Data	678
Decision	508
People	317
Business	251
Need	219
Want	163
Time	154
Analytics	158
Different	126
Customer	184

5.3.2 Comparison of word frequency with coding

In this section an analysis is conducted to determine to what extent the high frequency word groups (Table 5.3.1.1) were captured in the interview coding. This comparison of word count to codes, highlights the extent to which words indicated as important to participants through its high word frequency as per Table 5.3.1.1 were also captured by the researcher as important during the coding. Therefore, this exercise offers a form of assurance on the appropriateness of the coding process. This was done mainly for the top five word groups as included in Table 5.3.1.1. The analysis of each of the first five word groups are presented separately and each analysis reflects the codes used as well as the highest incidences of code co-occurrence, which is a count of the number of times two codes are allocated to the same or overlapping phrases, sentences or paragraphs of text contained in the interview transcripts. It is important to note that codes presented here are included on a judgemental basis and do not necessarily correlate to the thematic analysis presented later in Chapter 5.

Data

Given the field of enquiry of this study, it is not surprising that the word "data" was the most used word, as reflected in Table 5.3.1.1. The table below, table 5.3.2.1 includes the codes used during the coding process that include the word "data", and indicates the number of times those codes were used in brackets.

Table 5.3.2.1: Codes related to data

Data	
Data quality (26)	Responsibility for data (7)
Data sharing (22)	Data structure: access to data (5)
Extent of data use (18)	Ethical use of data (4)
Data used: structured vs unstructured (17)	Data: understand and leverage (3)
External data (13)	Overcome data quality challenge (2)
Use of qualitative data (9)	

As illustrated by the codes in Table 5.3.2.1 leaders offered their perspective on data quality and sharing or access, often voicing concerns around data quality and completeness as well as approaches implemented to address the challenges in data quality. The application of external data to augment internal data sources also received some attention, often as a future ambition rather than a current practice. Similarly, the use of structured data dominated, with aspirations to augment the structured with unstructured or qualitative data also dominant. The ethical use of data was also raised, but was not widely echoed by participants.

The code co-occurrence in ATLAS.ti, also reflected the fact data sharing and data quality co-occurred six times, as did data sharing and "get the basics right". This signalled the level of airtime given by participants to the need to have foundational aspects sorted out, such as having reliable data in place. Infrastructure and data sharing co-occurred four times, often due to participants noting the need to have the right infrastructure in place to facilitate sharing.

Decision [word group: think, thinking, thinks, decision, decided, decides]

Words in the decision word group were used when discussing an organisation's processes and patterns of thinking and delivering issues in making decisions. The codes that were used are presented below

Table 5.3.2.2 Codes related to decision making

Decision making	
Asking questions (21)	Business leading conversations (2)
Decision making enabler (14)	Decision making expert (3)
Decision making: empowering staff (3)	Business practitioners past decision making (2)
Analytics to replace expert knowledge/judgement (3)	

Table 5.3.2.2 reflects that the word decision as shown as important in Table 5.3.1.1 was often used by participants in the context of decision making. The coding indicated that big data and analytics was viewed as a decision making enabler, and that decision making included a process of questioning to ensure big data and analytics are applied to the appropriate problem. The codes "decision making: asking questions" and "collaboration" co-occurred four times reflecting also the overlap with a collaborative decision making process. Participants discussed the replacement of experts as decision makers by automated decision making processes, as well as the use of trusted analytics to augment decisions that were previously made based on expert knowledge.

People [word group: people, people's, peoples, person]

The third major word group related to people. The codes used that related to people are presented in Table 5.3.2.3 below.

Table 5.3.2.3 Codes related to people

People	
Collaboration (48)	Communication (8)
Skill sets (42)	Overcome data scientist/IT skills challenge (4)
Leadership support (29)	Lack of understanding by top management (3)
Adoption (22)	Team structure (1)

People	
Change management (18)	

The codes related to people-dynamics was discussed within the context of staff and the manner in which they work together as teams, and often in more diverse and cross functional teams from that which dominated before big data and analytics became more prevalent. The skills required also featured strongly, as did discussion around the support for the big data and analytics journey offered by leadership. Similarly, the initial resistance to becoming data-led, and means of overcoming this resistance through change management, was also referenced. The code co-occurrence analysis showed that the codes adoption and collaboration occurred together five times. Change management and collaboration occurred together four times, as did adoption and change management. This was indicative of the changed dynamics of how people work together and how they needed to be managed to get adoption of the right behaviours.

Business [word group: business, businesses]

The word "business" was often used in the context of a reference to the organisation being discussed, and an argument can be made that it should have been omitted as a common word in the same way that references to "I" or "you" were omitted. However, the word was also used in describing the business and how it operates to create value from big data and analytics, hence codes related to organisational aspects are referenced in Table 5.3.2.4.

Table 5.3.2.4 Codes related to business' organisational structure

Business
Organisational structure: centralised (13)
Organisational structure: decentralise (18)
Org size: speed of change (3)

The organisational structure and how this facilitated or hindered the data and analytics journey received a fair amount of attention in the interviews, and even though less general, observations around the time it took for organisations to undergo any necessary changes were also referenced.

“Change management” and “business scepticism” co-occurred three times and was indicative of how businesses approached change. The fact that “organisational structure: decentralised” and “collaboration” co-occurred five times indicated the extent to which leaders viewed the manner in which a business was structured as influencing the people-dynamic as well. “Organisational structure” and “data sharing” co-occurred four times as did “organisational structure: decentralised” and “infrastructure” as leaders shed light on the infrastructure impacts of organisation structure.

Need [word group: need, needs, needed, needing]

“Need” from a business perspective was often used in the context of customers (the tenth most frequent word group per Table 5.3.1.1). The identification and satisfaction of customer needs were cited as an increasing focus area where businesses has used big data and analytics. For example, RP6 noted that: “The way that we do that would be through modelling, so that’s one aspect, so how do we build propensity type models, to either increase take up, so for example give me all the guys who qualify for credit, and just because he qualified doesn’t mean you **need** it”. It was also used to express a business need, but often as it relates to customer. For example: RP7 stated that “in our [insurance business] world you **need** actually lots of data to predict the probability of individuals are actually going to die”. These quotes not only illustrated that needs related to customers, but also that customers were referenced as the source of revenue or value.

Table 5.3.2.5 Codes related to needs.

Needs	
Improved customer experience (16)	Empower customer with analytics (3)
Single view of customer (8)	Entrench customer (2)
Change customer behaviour: rewards (8)	Customer lifetime value (1)

As illustrated in Table 5.3.2.5, participants generally discussed the fact that big data and analytics enabled an increased focus on customer experience. They also referenced the need to better understand customers to be able to achieve the better experience. Similarly, participants discussed the need to change customer behaviour through incentives to achieve behaviours that were more value generative.

5.4 Thematic analysis

In this section, an analysis is structured to provide an understanding of how the data related back to each research proposition. Themes are analysed to determine the extent to which it supported the propositions put forth in this research (Chapter 3). Themes represent a judgemental grouping of codes that are related to one another. Even though the thematic groupings indicated\ a discreteness in the codes, the researcher’s actual experience was that these codes are often difficult to distinguish and group separately under themes. The researcher exercised judgement, informed by the literature review, in deriving the code grouping into themes as outlined in the analysis. Thus, codes were grouped into themes that were tied back to related research propositions to offer support thereof.

For ease, the research propositions are repeated first and then followed by a discussion of thematic findings from the interviews. The researcher kicks off each discussion with a table, presenting the granular codes that have been grouped into the theme being discussed. This is followed by a narrative overview of the coding process followed by the researcher, noting some of the thinking applied.

Thereafter, the analysis is done, first by providing a quantitative overview of the findings. The quantitative overview was prepared using “code document table”, an analysis tool in ATLAS.ti, to produce a table containing information on the extent of theme discussion. Each table is arranged to present four rows- a quote count, word count, total word count and relative percentage of speech. This table is ordered with participants represented in the columns, and columns are ordered based on highest relative percentage of speech- showing the participant who discussed a theme the most, first. The quote count and word counts and relative percentage lines represent the following information:

1. A quote count was undertaken to determine the number of quotes related to a specific theme that was made by each participant.

2. The number of words spoken by each participant, in relation to each theme, was counted and presented.

3. This absolute word count was converted to a percentage of the total word count per interview to determine the relative percentage of time participants dedicated to discussing the theme. The relative word count percentage is indicative of the amount of time a participant spent speaking about that concept or topic and therefore the importance they attach to that topic. It is important to note that attempting to add the percentages per proposition across all participants will result in the a reflection that

words spoken exceeds 100% of what was actually said- this distortion is due to code co-occurrence

Finally, to add a “thick description” of the findings, the researcher has judgementally extracted select quotes that best articulates a summarised version of the participants’ views. Quotes are included as indented paragraphs, in italic font. Emphasis is selectively added by the researcher to highlight the important parts of speech – emphasis is indicated in **bold** text formatting. Any text inserted to clarify parts of speech is indicated through the use of square brackets: “[ ]”. Quotes are generally presented to reflect the order in which participants gave it airtime, with cross references made between participants where they pick up on similar or contrasting points. Furthermore, in some instances the order in which quotes of a single participant are presented differed from the order in which it arose during the interview. This was done in instances where the researcher thought that such an alteration did not change the nature or meaning of the quotes, but the reordering provided clarity and emphasis on valuable insights.

The analysis of each theme ends with a conclusion in which the findings are summarised so as to pull together the various pieces of evidence presented in the research findings. Detailed discussion of the findings and their meanings are however not presented, as this is captured in Chapter 6 of this research paper.

5.4.1 Research proposition 1: Decision making

The first research proposition, as presented in Chapter 3, centres around the leaders’ perspective on the use of data and analytics in decision making. The proposition suggests that data-led decision making will be embedded in an organisation only if leadership demonstrates support for it. Furthermore, it puts forth that data-led decision making requires a different decision making process, in which asking the right questions and collaboration become more important.

To create value from big data and analytics:

- Organisations need **leadership** that **support** the use of big data and analytics in decision making;
- Organisations need to have the ability to **ask the right questions** at the right time;
- Organisations need to have collaboration between their data and analytics functions and business.

The three aspects of decision making per the above research proposition, are analysed and discussed separately in the sections that follow.

5.4.1.1 Leadership support for data-led decision making

The relevant part of the research proposition which is analysed in this section, is re-presented here, for convenience:

To create value from big data and analytics:

- Organisations need **leadership** that **support** the use of big data and analytics in decision making;

Coding for leadership support of data-led decision making

The researcher used the following granular codes related to decision making as a theme:

Table 5.4.1.1.1: Codes related to decision making

Decision making		
Lack of understanding by top management (3)	Business scepticism around D&A (4)	Org size: speed of change (6)
Leadership support (29)	Decision making: empowering staff (3)	Relevance and value (5)
Overcoming lack of understanding by top management (4)	Decision making enabler(14)	Time horizon (9)
Business practitioner past decision making (2)	Decision making expert (3)	Trusted analytic (7)
Extent of data use (18)	Expectations of D&A(4)	Use of analytics (12)
	Prioritising (20)	

The above codes were discovered during the process of coding for decision making, and were informed by the initial coding table. In the coding process, the researcher looked for reference to the ways in which decisions were made – both before and after the use of big data and analytics. Improvement or changes in decision making were therefore also given attention, as did references to making use of data, intuition, or judgement in support of a decision. The views expressed by leaders on the decision making process was also noted. References to the amount of time used in making decisions and how prioritisation was done in a world where more things are possible due to big data and analytics were also coded. The evolving nature of expectations of what data and analytics will enable organisations to do was also captured.

Overview

Table 5.4.1.1.2 reflects the extent to which decision making themes were discussed by each participant. The table is ordered in descending order of occurrence.

Table 5.4.1.1.2 Decision making – themes per participant

Decision making	RP4	RP6	RP2	RP9	RP3	RP5	RP7	RP8	RP1	Total
Quote count	31	17	18	15	11	14	11	7	12	93
Word count	2685	2543	1867	1900	1230	1358	979	599	744	13905
Total words	7465	7542	5737	8935	6142	7848	8319	6940	7956	66884
Relative %	36%	34%	33%	21%	20%	17%	12%	9%	9%	21%

In line with the underlying codes, the above table is reflective of the extent to which participants discussed aspects of the decision making process and the extent to which big data and analytics are used in decision making process itself. Furthermore, it also captures the views expressed around the support demonstrated by leaders for the use of big data and analytics within the context of decision making.

Noteworthy findings was that the overwhelming majority of participants did not view their organisations as using big data and analytics extensively in decision making, although it was utilised to a limited extent. This is despite the fact that five of the nine participants noted that their leadership was supportive of the big data and analytics journey. A lack of proven results and scepticism, long delivery time, a lack of clear strategic direction, and funding challenges were also discussed by participants.

Analysis of transcripts

In discussing the extent to which big data and analytics was used to make decisions, RP4 provided a view that data was used to inform decisions, but that it's application was fragmented, rather than pervasive.

> *"So, the **ability to imbed data lead thinking** and strategic thinking from a data perspective, end-to-end, **is lacking** and I think that the reason that it's lacking is **because it's not necessarily being driven from the right level by organisations.**"*

RP4 further noted that this lack of data-led decision making was linked to a lack of clear strategic guidance from an executive level to exactly how data should be used, arguably leaving business to discern this for themselves.

> *"I have* ***not come across*** *yet, call it a* ***'formal way' in which data decisions are made*** *within the organisations.* ***Organisations have a data culture,*** *so to speak, that is* ***infused****, but I* ***haven't necessarily seen anything formal around coming out of an executive layer*** *to say 'this is how we are going to use data end-to-end."*

Offering a hint to a lack of strategic leadership, RP4 explained that a knowledge gap existed at a senior level as data conversations were technical and not relayed back to its relevance in business operations, such as in decision making. Only once this narrative was changed would leaders demonstrate deeper buy-in to its application in business.

> *"The* ***data*** *conversation is largely misunderstood, especially* ***in the senior part of the organisation*** *and the reason it is misunderstood is because it* ***often becomes a technical conversation as opposed to a business conversation*** *and if organisations are going to make this data thing work for them... simplifying that [data] story and* ***then getting the boards, executives and levels below that to buy-in on the back of that****."*

RP4 believed that this was slowly changing as data and analytics required different methodologies, which pulled the executives closer to the big data and analytic projects to connect with all levels of business and understand how to use data better.

> *"Executives are slowly starting to realise that they need to understand, and* ***understand at a lower level of detail, what the organisation looks like from a data perspective,*** *how data can really be used to inform and enable a strategic imperative and cannot necessarily dictate sitting at the top as they traditionally used to do...* ***it's more regular thinking around****: how am I going to use my data better,* ***how am I going to use data to inform better decision making****."*

Given the size of financial institutions, RP4 did not believe that a change was likely to occur quickly.

> *"And I've seen in a number of banks and insurers to say the least that* ***it's not an overnight change*** *to get organisations as big as our tier one banks and insurers in the country to turn around will* ***take 18 to 24 months from a change management point of view****, so it is not going to be an overnight switchover."*

In discussing the extent of data use, RP6 noted that even though data existed, there seemed to be a lack of clarity on what to do with the data. RP6 therefore agreed with RP4 - that an absence of strategic direction for the use of big data and analytics existed, which the researcher understood had created a barrier to its widespread use.

> *"So I think in terms of bank as whole,* ***we're not short of data*** *however we are short of what we want to do with it. As an organisation* ***we need to develop much [more] strongly on the strategic purpose*** *of data and analytics."*

RP6 noted that it, in certain cases, scepticism had also created barriers in getting the organisation to use big data and analytics. These mind-sets existed on the part of the business, who felt that part of their function was made redundant by the data and analytics team. He referred to the need for a "segment" or "organisation" approach to change this. The researcher interpreted this as a need for top-down support to be demonstrated to dispel the feeling of threat and clarify the empowering nature of having data available to better aid business practitioners.

> *"It goes to this thing of,* ***how do we enhance the roles or the efficiencies of our staff without making it seem as if we try to take away what they do****? So that again, it goes* ***to either a segment approach or an organisation approach****, by saying 'Guys,* ***we want to empower you with the information,*** *but we don't want to dictate to you with what you should do with it'".*

RP6 also noted that decision makers had a mind-set considered the analytics team as working for them, rather than with them in the decision making process, and that this was not conducive to the use of data and analytics in decision making. RP6 believed that the data and analytics teams needed to build a brand and a reputation for themselves to create the required buy-in the contribution they offered.

> *"So the challenges are that:* ***how dare you say no, because you are meant to be working for me type of thing?...*** *So it's not being the loud voice type of thing actually shoots you in the foot, that goes to* ***building brand, reputation****."*

In discussing the extent of data usage, RP2 acknowledged that people still grappled with challenges that the analytics team had not solved through big data and analytics, and instead they relied on other means to inform decisions, indicating that big data and analytics are not pervasively used in his organisation.

> *"But you also have, as we are lucky to have in foreign exchange, a lot of people that are looking after non-quantitative functions, that have **a lot of experience, good instinct, gut feel, understanding, appreciation for the business, that they are not always able to express quantitatively.** It is critical for us to understand what kinds of questions that they grapple with, and **then convert that into some sort of quantitative, actionable, formulaic, representation**."*

RP2 also acknowledged that a lack of understanding of big data and analytics sometimes meant that there were unrealistic expectations of the extent to which it can be of use in business. This is consistent with a lack of understanding noted by RP4, although the nature is slightly different.

> *"We have spoken about this [big data and analytics] **industry still evolving over time.** It is not a static thing, and people kind of get glimpses into that, through a talk here or TedTalk or conference that they have gone to, and **often it is portrayed in a manner in which it is going to solve all your problems in a night**. So they have got a **misunderstanding of what a data scientist does**, and then they have got **lofty expectations** which are not aligned with reality. So you have very annoying and **unrealistic expectations**…"*

Furthermore, RP2 mentioned that the ability to generate insights to inform decision making was negatively impacted by the time it took for big data projects to reach fruition. Time delays were due to the time it took to get priority in a sizeable organisation, with competing objectives. This resulted in ideas often not being implemented. Therefore, even though data-led decision making may be desired, it was not always possible within the organisational construct.

> *"So then there is, the challenges in my team, of working in a team that is 600 strong, but part of a 40 000 strong organisation. So, all these fantastic ideas, how do they see the light of day? If you require system implementation, or development, **you have to get priority. That is incredibly difficult** to achieve, and along with that, comes issues of **staleness**, so you have found this gap, but the **window of opportunity closes before you are able to implement.** And, that is hugely frustrating to younger staff members, whose ideas just die before they have been given a chance."*

Despite these issues, RP2 explained that to start overcoming these barriers and add value, the work undertaken by a big data and analytics team needed to be viewed as valuable by executives. Fortunately, this was the case within his organisation.

> *"You have the challenge of* ***ensuring that the work that you do is relevant and that it is creating value and that it is perceived by the executives****. You need an environment that is supportive of that kind of work.* ***Thankfully, we have that in our space****."*

RP9's organisation has not always used data analytics, but due to leadership buy-in to becoming data-led, this had recently started to change.

> "*So, I think culturally, as I said,* ***we've never really been a data or a process organisation****, but I think we are there now and as I said,* ***senior management buy into data absolutely****."*

RP9 expanded on the leadership support demonstrated by the CEO of their organisation who actively considered how data and analytics could be applied to solve problems, and was leading the way in shaping the belief in the power of information.

> *"Our new CEO took over two years ago, let's just say, and in the last year* ***he's really been driving the analytics agenda*** *and saying, 'How do we use…' So, he's really pushing the analytics agenda.* ***He's now got an analytics team****. It's small, it's in its infancy, but it's growing* ***believing [in] the power of information****."*

At the same time, RP9 elaborated to explain that although buy-in existed at the executive level, the CEO had a misconception around the operational maturity of the organisation and did not understand the amount of effort involved in producing insight, and this may result in reluctance to invest in creating the right capabilities.

> *"He [CEO]* ***doesn't really understand what's going*** *on in this place, so when he asks for info* ***people scramble to pull bits together for him****,* ***taking hours*** *and hours and hours, but* ***he gets the final result*** *and he goes, 'Oh great, it's all there,* ***there's no problem', but actually there is a problem. So, if we do need more money at some point…*** *he will not understand why we need it because right now his attitude is like*

> *'**everything's fine, I get what I want, why would I need to spend more money on this?'** But **he doesn't understand** the scrambling that goes on to get him anything."*

RP3 noted that data was not used extensively within the organisation.

> *"At the moment I think it's **very small section of the business that sort of makes fully data-informed decisions.**"*

RP3 explained that their organisation was in the early phases of the journey and consistent with RP6, that there was some scepticism to be overcome around the value of big data and analytics and thus RP3 initially experienced funding challenges.

> *"So it clearly hasn't fully changed, **it's obviously a journey and we're in the early days of it**, but **initially** the problem was when we try to prove the possible value-add of the embarking on this plan and strategy for the bank, there was **scepticism** because **we required some money** and some funding and it was **what will be the return of value**?"*

RP3 expressed a view that the executive, strategic layer of the organisation supported the big data and analytics initiative, however RP3's view was that at an operational level buy-in did not exist. The researcher understood this to be part of the reason that there was a shortfall in the use of big data and analytics throughout the organisation.

> *"So especially **at the senior leadership, there's lots of support** for innovation at the **data counsel** and forums like that. **Definitely a port into the strategy, but people on the ground we need to interlock with it to land the value**, to implement the new insights of new data products, to synchronise into the business areas… So **from the top is where there is buy-in, it's the guys on the ground that you need to actually build with that requires a bit more effort.**"*

As a potential reason for the lack of operational buy-in RP3 concurred with the view expressed by RP2 and noted that big data projects took time to implement. This resulted in the risk of businesses pursuing another solution. To combat this, RP3's team have tried to keep business well-informed of progress to create understanding and patience.

> *"It **takes a long time** … it easily takes **six months to a year** at the moment depending on the tool. I mean there could be **business getting a bit of cold feet** but **as long as***

> ***we include them** in the journey the whole way through, keep them updated on a bi-weekly basis….”*

RP5 believed that big data and analytics was used extensively within their organisation and that it enjoyed significant support by their CEO and leadership team.

> *“**I think that it is completely transformational…** big data and banking. I think the digital transformation has already happened… I think from a **top management perspective I think we have at least strong buy-in to continue with the journey**.”*

RP5 explained that historically decisions, such as credit decisions, were made on a judgemental basis, whereas big data and analytics had enabled decision making through modelling of outcomes. This change was motivated by the need to improve the quality of decisions, illustrating the faith the organisation placed in data-led decision making over other methods.

> *“..the vast bulk of credit decisions in the bank are done fully automated with no human intervention. .. the big change **over the last decade** was **moving from judgemental decisions making to automated decision makin**g, **the big driver of that,** ironically, **was not reduction or cost or speed… but quality of decision making**.”*

RP5 stated that the transition had not been easy as people were reluctant to trust machines to take decisions, indicating scepticism with the changes enabled by big data and analytics, similar to RP6 and RP3.

> *“It’s always interesting to find how **people struggle to trust the machine** to do something, even if the **machine can do it a lot better**.”*

RP5 also picked up on the timing concern raised by RP2 and RP3 and stated that practical trade-offs needed to be made to enable the organisation to use big data and analytics - what is possible needed to be balanced with what could be delivered in the short term.

> *“We have to then quite **be deliberate** if we introduce new variables that we want to use in decision making process... when we will be able to introduce better decision making approaches based on those. There are **quite a lot of trade-offs between what is possible** in the laboratory in the data side and between **what actually we can practically deploy in short and immediate term**.”*

RP7 believed that big data and analytics was not used optimally to make decisions in his insurance organisation.

> *"If I had to be honest, I'd say no.* ***I don't think we've done enough with data in our decision making capabilities****."*

RP7 explained that due to limited successfully implemented big data and analytics initiatives, there was a lack of buy-in, given that the evidence of the value big data and analytics could bring was scant. RP7 noted that there was a theoretical appreciation for the potential of big data and analytics and what it could bring to the organisation.

> *"We* ***haven't had lots of successes*** *we can sort of celebrate and say we've done the following and this is translated into the following client behaviour, the following improvement in cross selling and therefore the following improvement in the bottom line.* ***I don't think we're there yet*** *and there's is an appreciation I think we can do all of these things. I think people are appreciative definitely to say that they believe that* ***that's the future and to go there****."*

RP7 elaborated that a potential reason big data and analytics was not used extensively was because projects did not get realised in the manner envisioned in the original business case and noted that projects that do not result in models that can be used to support data-led decision making should be abandoned. The researcher found the sentiment of this comment resonated with the views raised by RP2, RP3 and RP5's statements about how long implementation timeframes made big data projects redundant. This implied the need to monitor the viability of big data projects.

> *"So our discipline around putting a* ***business case*** *on the table, actually* ***tracking*** *it throughout the journey…* ***Be willing to pull the plug*** *if something is not working!"*

RP1 noted that that extent of use of data and analytics varied in his organisation, but that the leadership did buy in to the use of big data and analytics. Based on the general impression gathered from the interview, the researcher noted that areas of the organisation used big data and analytics, but the use was not consistent or pervasive.

> *"It* ***varies****, is the short answer… There's that leadership to say: 'We need to do this, so we will invest in the technology'."*

RP8 explained that big data analytics was used within his organisation, but not extensively. RP8 believed that there was more that could be done with the data and that management needed to move toward being willing to invest without expecting immediate return on that investment.

> "*If you're asking me 'How is this organisation doing* ***in terms of data and analytics?' I think this business being mainly about managing risks that's…****you can still build a lot of insights from the data that we have. But also, I think just* ***leadership saying: 'okay, we are going to have the courage to spend money and not necessarily expect immediate return'****. Mining from the insights is just one of those things."*

Conclusion

The below table provides a summary of the key points noted from the analysis of decision making.

Table 5.4.1.1.3: Summary of findings: leadership support of data in decision making

Decision making	RP4	RP6	RP2	RP9	RP3	RP5	RP7	RP1	RP8	%
Data used pervasively in decision making						X				11%
Not pervasive/limited data utilisation in decision making	X	X	X	X	X		X	X	X	89%
Limited or no leadership support/buy-in	X	X					X		X	44%
Adequate leadership support/buy-in exists			X	X	X	X		X		56%

As summarised in Table 5.4.1.1.3, eight of the nine participants interviewed during this research expressed a view that data was not used pervasively for making decisions within their organisations. Only RP5 noted that his organisation used data extensively. five participants believed that data and analytics received support from the leaders of the organisation, while the remaining four participants did not expressly note strong support. This was interpreted to mean that limited or no leadership support existed. These findings support a conclusion that data-led decision making is applied, although not pervasively and although the majority of leaders support big data, a lack of sufficiently strong and consistent leadership exists.

In addition, it emerged from the analysis that RP4, RP2 and RP9 believed that a lack of understanding exists on the part of their organisation's leaders. Funding was raised by RP3, RP9 and RP8, indicating that its consideration is important and that it could present a potential limitation of wide-spread adoption of big data and analytics in decision making. RP3, RP2, RP5 noted that the time it takes to implement a big data and analytics initiative had negative impacts on its adoption in decision making, and RP7 noted that viability assessments had to be undertaken to assist with managing this. RP3, RP6, and RP5 explained that their organisations faced scepticism with regard to what big data and analytics could enable them to do. RP4 and RP6 also noted that clear strategic direction around how to leverage big data and analytics was lacking.

5.4.1.2 Asking the right questions

The relevant part of the research proposition which is analysed in this section, is re-presented here, for convenience:

To create value from big data and analytics:

- Organisations need to have the ability to **ask the right questions** at the right time;

Coding for asking the right questions

Based on the literature reviewed, the researcher created a code relevant to this theme. In this case, only one code was used, which is asking questions, as per Table 5.4.1.2.1.

Table 5.4.1.2.1 Granular codes related to asking the right questions

Asking questions
Asking questions (20)

Part of the decision making process that involves formulating an appropriate question to explore using big data and analytics is important in creating value. During the coding process, the researcher flagged instances where reference was made to the process of asking questions as part of the decision making.

Overview

The table below summarises the extent to which participants discussed the process of asking questions during decision making.

Table: 5.4.1.2.2: Asking the right questions - themes per participant.

Asking question	RP2	RP6	RP8	RP1	RP4	RP3	RP7	RP5	RP9	Total
Quote count	4	5	5	3	2	1	1	0	0	7
Word count	708	710	411	298	230	82	87	0	0	2526
Total words	5737	7542	6940	7956	7465	6142	8319	7848	8935	66884
Relative %	12%	9%	6%	4%	3%	1%	1%	0%	0%	1%

As indicated in Table 5.4.1.2.2, the process through which questions are raised, refined and are tackled through big data and analytics programs was discussed by seven participants, with two executives (RP5 and RP9) not offering any thoughts on this topic. The level of airtime given to the topic of asking questions was relatively low across all participants.

Noteworthy findings was that participants believed that the process of asking questions during decision making had changed, and the majority of participants believed that the analytics teams could take the lead in formulating the right questions to identify business problems and discover novel insights. However, the thinness of the responses led the researcher to conclude that the process of asking questions had changed, there was not much information on making sure the right questions was asked as a key part of decision making, enabled by big data and analytics, from the perspective of the leaders participating in this study.

Analysis of transcripts

RP2 noted that often business may have questions that the big data and analytics team historically were not equipped to address and this had resulted in a form of disillusionment. The analytics team therefore have to make the effort to unmask these business problems that practitioners may be grappling with and acknowledge them for having originated these.

> *"A lot of these [business] people **have become jaded over time** because they had these questions but no one could answer if for them, so it was sitting there, kind of blatant. So it is important as a **technical team to get close to business, to try and understand the business as much as possible,** not operate an ivory tower; give people credence for the value that they bring"… "It is critical for us to **understand what kinds of questions that they [business] grapple with**, and then **convert that into some sort of quantitative, actionable, formulaic, representation**."*

RP2 elaborated and explained that questions that originate in business have to be pursued as a partnership by business and the analytics team, if the true value in that question is to be unlocked. The process could not be in the form of a set of instructions issued to the analytics team, or the analytics team working on their own.

> *"Ultimately, what is going to emerge is…**the value of it is going to be how well you have answered someone's questions** that they have conceived of, and how you have t**aken them along the journey of exploring different avenues** as a data scientist… It's **never** going to be anything **optimal if the questions are a set of instructions that emanate from the business person or something that you are relied upon to do in isolation as a data scientist**."*

Furthermore, RP2 stated that the analytics team also had a key role to play in seeding new ideas and that this role was often unfulfilled due to the analytics team's time being spent fulfilling requests for information that were provided by business, rather than experimenting to see what new information can be discovered. RP2 overcame this by setting up a specific team, the insights team, who was mandated to explore and find insights that can used as the catalysts for business decisions.

> *"I think that a second dimension for me is whether you are able to **provide novel, useful and actionable insights** into the business. Particularly, where the questions don't emerge from the business. **It is helping them understand, or bringing to their attention something that they have no knowledge of, or that they inkling about.** I don't think that can be underestimated. And **I don't think often you have the latitude as a data science team to be able to do that,** because more often than not, you are following instructions, **which is why the insights team is so important because they are experimental.**"*

RP6 expressed support for cooperation in formulating questions, and agreed with RP2 that either business or the analytics team can raise a question. RP6 went further and noted that engagement between the parties to refine the question was key; the analytics team could not be in a position where it simply does what business asks them to. This suggested a form of autonomy for the analytics team in looking for novel ideas from using big data and analytics. In this regard, he was therefore also in agreement with RP2.

> *"Anything that we do as **business as usual**, quite easy. It's a thing of saying the idea can be **brought up by either us or another area** and the **next stage** as I call it, is now that you have an idea, let's us actually go and brainstorm what the idea is about. That*

> *idea then **needs to be critically evaluated** to say: really, is it actually that bad a problem, or not really so much? That in between stage of assessing whether the idea is relevant or not, is very important. **It's not just a thing of saying, because your stakeholder told you to work on something, therefore you work on something**."*

RP8 echoed RP6's comment and described the data scientist's role in asking questions as asking the "right questions", which are questions that business would not ask. Through this process a level of challenge was offered to assumed practices and fresh insights can thus be generated. In this, the researcher found that RP8 appeared to be in agreement with RP2's assertion that the role of the analytics team is finding new, novel information, rather than tackling normal business questions.

> "*Your **data scientist** won't work on "how many cars did we sell?" and "how many did we convert", "how many did we approve?" and "what do our vintages look like?". It's not about that, **it's about generating the right questions**."*

RP3 viewed business as having many ideas that they cannot express and require collaboration from analytics teams to articulate the questions properly. This picked up on the point raised by RP2 and RP1 around business requiring assistance to tease out and articulate a problem. RP3 pinned this down to a hesitation and an education gap on the part of business.

> *"However I think there's still sort of a challenge for the organisation themselves. They are thinking of **certain things they think will be cool for our team, they don't know how to use the lingo as well back to us;** so they don't know how to express themselves so if they explain the problem then we can say 'Okay. What you actually asking for is this?' So there's still some **hesitation and education on their side to understand the big data**."*

In contrast to the views expressed by RP2 and RP6, RP1 expressed view the analytics team could take the lead in asking questions, RP1 opined that the problems data and analytics should be applied to should originate from the business and can then be worked on in collaboration the analyst team.

> *"The **business owner needs to articulate the problem**. So not saying that the quant can't come up with the idea but that [business owner's idea] is what you start with; what's the business problem I am solving? What is the business opportunity for me to make more money? What's the business opportunity for me to lose less money."*

RP4 also agreed with other participants that the co-creation of questions was important. He explained that analytics teams could not go at the big data and analytics journey in isolation, they need to work with business which has the business problem to solve. He believed that structure of having an innovation hub leading the data conversation was counterproductive to this and noted that these areas needed to be subservient to business, which appeared in contrast with the view of not simply following instructions as voiced by RP2, RP6 and RP3 and was therefore more aligned with RP1's thinking.

> *"Many of the insurers and banks have created innovation hubs, have created **more standalone capability through which they are trying to drive, call it 'next gen data conversation'**… Innovation capability only **exists because we have a lot a of business problems to solve**. They can't do this in isolation… The innovation hubs have this hot-headed mentality in terms of that they were bigger and better than everyone else **but the way to achieve success is almost being subservient to the organisation**."*

RP7 also raised the importance of taking business on the journey with the analytics team in answering a question in a collaborative manner.

> *"I think our challenge is to, partly because of our operating model as well, it's taking these things and making it practical for our employees. **Our employees** have got a **very good understanding of, if we've got to be successful you need this, this and this and that's going to be the outcome**. And **take them continuously along that journey with proper communication**, so you can actually go and prove to the staff: you have now invested a lot of your development time to actually do this."*

Conclusion

Table 5.4.1.2.3: Summary of findings: asking the right questions

Ask the right questions	RP2	RP6	RP8	RP1	RP4	RP3	RP7	RP5	RP9	%
Changed approach	X	X	X		X	X	X			86%
Unchanged approach				X						14%
D&A team must have latitude to explore and seed novel ideas	X	X	X	-X	-X	X	X	N/A	N/A	71%

As illustrated in Table 5.4.1.2.3 the research participants generally believed that there was a change in decision making as it relates to the process of asking of questions. Interestingly, the view was not just in how and where in the organisations questions were seeded, but also in how problems were tackled - questioning was seen as an ongoing activity. Two participants (RP1 and RP4) believed that business practitioners had to take ownership to the process of asking questions, due to the depth of understanding offered by proximity to the problems- their contrasting view is indicated in a –X in the above table. They did not see it as optimal if the analytics teams embark on the big data and analytics journey in isolation; they needed to work with business that has the business problem to solve. Five participants expressed support for the data and analytics teams to have the scope to originate and explore question and, through this reach novel insights. RP6 and RP2 agreed that even if an idea is seeded in business, the analytics team should not get stuck in a paradigm whereby it simply takes instructions from business and does not contribute their skills to the formation and refinement of a business problem through asking questions.

Although seven participants to the study spoke around the need to ask questions in a big data and analytics world, two participants (RP9 and RP5) did not participate in this topic at all.

5.4.1.3 Collaboration

The relevant portion of research proposition 1 is replicated below:

To create value from big data and analytics:

- Organisations need to have collaboration between their data and analytics functions and business.

Coding for collaboration

In coding for collaboration, the researcher identified all references made by participants to a working relationship between data scientists and other staff members. This included references to manners in which data scientists operated when working with business as well as the various responsibilities of data scientists and the persons they partner with.

Table 5.4.1.3.1: Codes related to collaboration.

Collaboration	
Business leading conversations (2)	Data sharing (22)

Collaboration: business, IT & data scientist (48)	Data structure: access to data (5)
Methodologies (agile) (3)	Organisational structure: centralised (13)
Overcome collaboration challenge (5)	Organisational structure: decentralised (18)
Responsibility for data (7)	

Overview

As illustrated in Table 5.4.1.3.2, participants spent a substantial part of the interviews discussing collaboration. RP6, RP3, and RP2 were the top three devotees to this topic.

Table 5.4.1.3.2: Collaboration themes per participant

Collaboration	RP6	RP3	RP2	RP4	RP9	RP5	RP7	RP1	RP8	Total
Quote count	21	16	9	14	9	11	8	14	3	84
Word count	3257	1479	1206	1420	1332	1048	927	660	395	8467
Total words	7542	6142	5737	7465	8935	7848	8319	7956	6940	59342
Relative %	39%	24%	21%	19%	15%	13%	12%	8%	6%	14%

The coding process revealed that participants discussed collaboration at three levels. Participants discussed collaboration between big data and analytics teams and the business areas within the organisation that they service. Participants also provided observations about collaboration between various data and analytics teams within the organisation. Finally, collaboration with external parties was also discussed.

All participants noted that the big data and analytics journey required greater collaboration and the general view was that this collaboration did not happen organically, but had to be encouraged and driven through initiatives within the organisation.

Analysis of transcripts

Findings around the collaboration between analytics teams and business areas are discussed first, followed by collaboration between analytics teams and external clients. Thereafter, the findings around collaboration between various data and analytics teams within the organisation are presented.

Collaboration between data analytics teams and business practitioners

RP2's organisation had decentralised big data and analytics teams. RP2 noted that big data and analytics teams were at the service of business stakeholders and that combining the technical skill sets with the business skill sets, through teamwork, enabled the analytics team to adapt a different perspective and understand the business problems they needed to solve. Through collaboration, business problems were also approached and solved in a different manner than in the past.

> *"So it is important as a **technical team to get close to business**… I would say that ultimately you are **serving the business** and it is important to understand what the needs are of these different stakeholders… **So you have got to bring these technical skills and the business skills, and it is that collaboration that is critical**… It is an area of consultation with the business **that allows them to think about things in a different manner to how they may have in the past**."*

RP2 noted collaboration is a mind-set or a culture that is not automatically in place, but needs to be developed if optimal results are to be achieved.

> *"That is the kind of **thing that you develop with a bit of practice and mind-set that is one of engagement** with the business partner, always. So, it's **never going to be anything optimal** if the questions are a set of instructions that emanate from the business person or something that you are relied upon to **do in isolation** as a data scientist. It has got to emerge…and **it's a culture that you develop**."*

According to RP3 collaboration had historically been lacking within his organisation.

> *"People have always been doing stuff but it was **never collaborated or stitched** in the same places together."*

RP3 also mentioned that a reason for this was that business practitioners did not understand the relevant terminology to express their problems to the analytics teams, and that this may have hindered the extent of collaboration. To overcome this, the analytics teams partnered with business practitioners and took them on the journey to show the connection between the work that was commissioned and that which was delivered to bridge the education gap and facilitate better collaboration.

> *"They are thinking of certain things they think will be cool for our team, **they don't know how to use the lingo as well back to us;** so they **don't know how to express***

> ***themselves*** *so if they explain the problem then we can say 'Okay. What you are actually asking for is this.' So there's still of* ***some hesitation and education on their side to understand the big data****…* ***We partner with them*** *and show them why did we do things like this and like that. So whatever problems they are solving falls onto the data analytics teams group of work, and* ***there's immediately a connect between what we're doing and what they want.****"*

RP7 agreed with RP4 and noted that within his organisation, there has been a deliberate attempt at fostering collaboration, particularly to overcome a concern that the big data and analytics team (referred to as the centre of excellence) could, through their work, make other functions redundant. He raised a powerful point: for collaboration to work, the data and analytics team need to have credibility and deliver value.

> *"We had* ***big projects on just focusing on the collaborative culture*** *that includes all the aspects of share success appropriately, work together, share information sort of on a better basis. And I* ***can't claim that we are there as yet****, but we've made, I think, excellent progress where* ***people start to understand that this is not going to take my job away****... But the* ***centre of excellence has got to prove to the previous guys*** *who were involved in these things that might now sit in the segment, that 'listen, we can deliver these things and we* ***can appropriately make sense of the figures that we can add value in your life****'."*

RP6 also mentioned that business had come to buy into the value that data analysts bring in gaining insights from data and, as a result, the analytics teams was more proactive in creating solutions.

> *"The other things that's also worked quite well was the* ***evolution of business buying into what we've done – so in other words data by itself useless, you need to have people that understand what to do with the data.****... They [business] explain the problem and* ***we come up with solutions for that problem, and we collaborate with them*** *in creating that solution."*

When it came to internal collaboration, RP9 also noted that silos existed within their organisation, and this had manifested in an inability to cross-sell various products offerings of the organisation to one consumer. This indicated a lack of collaboration.

> *"So historically, very siloed. So again, there are* ***quite distinct business units.*** *I think in the last year there has been a big drought, so we've got cross-sell opportunities that are just sitting there waiting and we're not taking them."*

RP9 spoke to the fact that their organisation's big data and analytics team was centralised, but not necessarily within the right executive reporting line to enable it to service the entire organisation, and this hindered broad-based collaboration.

> *"My argument is that that person, that team needs to report centrally to somebody like me...* ***So, I don't know what the [best reporting line] answer is****, but I just know you've got to keep the whole executive happy or you are not going to get buy-in into that structure [or] into analytics."*

RP1 noted that in the big data and analytics world, the ability of teams to collaborate had become more important as the modern way of working.

> *"We have to change it…* ***behaviours need to change****… so you don't have to own everything, but* ***your ability to collaborate and make it win-win... that's where the gold is****."*

RP4 noted that methodologies that have been adapted in the big data and analytics paradigm, such as agile, resulted in business being involved in a project from the design through implementation stage. This achieved greater collaboration as compared to other approaches in which business was only involved at the design phase.

> *"Because of that agile injection into the organisation that's* ***forcing business and IT to work together*** *because now our* ***business is not just part of a design conversation but actually part of the conversation from day one and they continue in that conversation****…* ***business starts owning the problem then they start owning what their data means to them****. Well, then it becomes a more* ***mature conversation,*** *as opposed to something which is we'll worry about data as an afterthought."*

RP4 also noted that it was important that a collaborative culture should deliberately be cultivated within the organisation to ensure people connected beyond their traditional work boundaries. That said, he noted that a change initiative was also more conducive to collaboration than an instruction from a boss to work together would be.

> *"The second one is very much **instil that collaborative culture**, call it a **formal change** management intervention… I have certainly seen an **uptick of more informal kind of events** that are organised so that **people start to understand each other outside of their traditional boundaries** and that is certainly **bearing fruit** in tow of the organisations that I am quite closely involved with… But I still think that **true in power collaboration** comes when people actually connect with people and not because my boss is telling me to do so."*

RP8 saw a lack of collaboration as a leading cause for the lack of traction in creating value from big data and analytics within his organisation.

> "*… **we have to get business to be at the leading edge of it**, but business has to be **exposed to see what's possible**… of all these five focus areas data is the one that's had the least traction."*

Collaboration between big data and analytics teams and the clients of an organisation

RP3 noted that their organisation also collaborated with its clients to empower them to improve their businesses, which in turn delivered results to their own organisation.

> *"When we combine data [with our clients'] and **we give increased insights to them** for their customers, they can use it to **improve their market share** which indirectly will come through to us because they bank more with us, which **makes more money for us. It's a win-win**."*

RP9 noted that legislation was driving an imperative to for insurance brokers to exchange data with their insurers, thus driving an enhanced form of collaboration.

> *"And now **legislation is forcing us** to just start doing that more and **insurers are saying that they want more access,** to even real-time data, because our business is all about risk."*

RP9 noted that a key enabler of collaboration with their clients was having integrated systems.

> *"Our other challenge from a leadership perspective, we have got this wonderful model planned in our heads **but integration with the insurer is a big issue**. We think we have got problems with data, they have probably got the same problems probably times ten and we see that when we are trying to do data exchanges, now even."*

Interestingly RP7, as an insurance institution, offered a counterpoint to the point made by RP9, as an insurance broker, around the sharing of information and insights produced by analytics teams. RP7 opined that information related to the end consumer of the insurance product should be shared between the two parties as a means of retaining both the relationship with the broker and end consumers

> *"And some of this for us is not necessarily client behaviour, **it will actually be put into an intermediary world as well**."*

Collaboration between different big data and analytics teams in an organisation

RP2 noted that the sharing of data between internal business units of the organisation was not well-developed.

> *"When it comes to internally-generated data from the product house, it is very easy. But there are many **opportunities to augment that data with other information that is available from other product houses** or let's call it things like statutory returns; there are things like cross-border payment monitoring that is done in the group corporate centre. **And none of these data have yet been appended to our existing data sets** or collated or injected into our data landscape."*

RP3 echoed this and noted that the disparate data sources used internally can produce different results when the same work is performed by different teams. In addition, he noted that the sharing of insights generated from data was also lacking.

> *"The **sources of data** people have used was **not the same** so you get **different results**, even if you had done the same work. No talking to each other. So there's no single view of data that is clean and conformed **and insights have never been shared, widely**."*

RP3 explained that his organisation has moved toward centralising the big data and analytics team to overcome some of these challenges.

> *"Hence you know the **shift of focus towards having to separate this function** and build it up in the firm, so it can be **deployed throughout the firm**. It's obviously a long term journey. So I think most firms and ours alike, have had imbedded analysts in the organisation through our time. People have always been doing stuff **but it was never collaborated or stitched in the same places** together…"*

RP3 believed there was active encouragement of collaboration, and that this is getting traction and realising efficiencies, despite the fact that the value was not always self-evident up front.

> *"So there's definitely **focus and encouragement to share across the group**… We've been collaborating and **sometimes it's hard for us to see the benefits** because we have the contacts, but once we get a lead we've got to present and vice versa; and sometimes we see similar things, it's just new data. So **we try and share, because of the move across the bank, towards a concept of wholesale banking, and trying to remove silo mentality and duplication**."*

RP5's bank historically had a decentralised approach with different areas in the bank viewing the data residing in that area as theirs, rather than belonging to the wider bank.

> *"We **traditionally had silos** where you maybe won't even believe it, for example finance and risk **didn't share information**… **People also felt that the data was theirs** and it didn't belong to the bank."*

According to RP5, an active drive was undertaken to change this mentality and this has had some success, although limited.

> *"So **breaking down these barriers deliberately** and getting people together in forums, that has been a v**ery big drive**... If I had to be honest, from finance and from a credit data perspective, I think there is quite good collaboration across the bank, **where there is very limited cooperation** let's say for example, would be **across operational data**."*

RP5 believed the approach had become more collaborative with the adoption of a somewhat mixed approach to collaboration across the teams in the bank. Teams try to solve the same problem individually and then get together to share ideas and knowledge gained in this manner. It implies a form of duplication, which may not be sustainable in the long term.

> *"So we've divided the data areas in the bank into domains so we get **analysts** that are working hopefully on similar problems or maybe say solving the same problem that **we have then working together, sharing data, sharing approaches** so we can get to better results and we ideally rather **have a smaller set of really powerful innovations or approaches** that we can then productionise better."*

RP5 expressed a view that this sharing of information was key to competitive advantage.

> *"There is a **significant competitive advantage from information across** different banking products, and even into areas like insurance."*

RP6 discussed a similar form of collaboration to RP5 where teams work individually but have access to information from a central shared depository.

> *"So we work siloed in the sense that **we know what we want to do with our stuff,** and it's also not siloed in the sense of where we access our stuff from, is centrally accessed."*

However, RP6's organisation had created forums within the organisation that assisted in creating knowledge sharing and therefore collaboration between different data and analytics teams.

> *"Yes, within the bank there is what they call a **Data ExCo,** but they talk about analytics as well, that's a once a month thing. You have various product houses, various channels, various segments and **pretty much people who are involved in analytics and data across the organisation,** so they come and attend that and basically it's a type of a little bit of s**haring ideas or of results**. …[if] there is a project that impacts all different areas **how do we as a bigger team go and dedicate people or time** in order to solve this bigger problem."*

RP7 explained their collaborative approach in terms of their organisational structure, and similar to RP6, had a two-pronged approach in which the IT and data are centrally located, while the actual work took place across various teams and functions.

> *"Our analytics has got a two-phased approach, we have the individuals that looks generically and they **need to touch base with some of the individuals that live either in the segments** or in the individuals that live in the centre of excellence and come up to find together sort of solutions. What we've done is **the actual IT component is together to manage, to collate the data,** but the **actual interaction and understanding and the analysis happens a little bit in different environments**…"*

RP7's organisation formed communities of practice so that different teams could build relationships, share interests and information and this has helped improve collaboration.

> *"Also sometimes, **you just have to create communities,** even though those communities might run across different reporting lines. So someone might sit in segment or client engagement solutions. If we can create communities, a lot of the time **I think it's about relationships**... You need to know these people, make them a community where they can **share interest**… So creating communities where you can **actually share information** better and I think that has helped us."*

RP1 expressed a view that big data and analytics required increased co-operation.

> *"So if you take business in general and it becomes more so in this area, **ten years ago 80% of what you needed to get your job done related to your individual skills** and 20% on your ability to collaborate, **that relationship is [now] completely inverted. Eighty percent of your ability to get something done is ability to collaborate** and only 20% of your specific skill."*

Conclusion

Table 5.4.1.3.3 provides a summary of some of the salient views expressed by participants.

Table 5.4.1.3.3: Summary of views expressed by interview participants

Collaboration	RP1	RP2	RP3	RP4	RP5	RP 6	RP7	RP8	RP9	%
Centralised structure			X	X				X	X	44%
Decentralised structure	X	X			X	X	X			56%
Organisation structure contributes to collaboration			X			X				22%
Organisation structure does not contribute to collaboration	X			X	X				X	44%
Collaboration is more important	X	X	X	X	X	X	X	X	X	100%
Collaboration outside the organisation is important			X				X		X	22%
Deliberate initiative to improve collaboration			X	X	X		X			44%

The participants of this study had consensus around the fact that operating in a big data and analytics world required a greater degree of collaboration, thus supporting this portion of Research Proposition 1. The findings suggested added dimensions to the construct of collaboration, in that organisational structure and change management was discussed by the participants as well. Support for collaboration existed irrespective of the organisational structure – i.e. whether the data and analytics function was centralised in the organisation or consisted of multiple decentralised teams. However, the manner in which organisations were structured had different implications for collaboration. The need for collaboration was not only noted with business areas, but also between the different big data and analytics teams in an organisation where these functions were decentralised. Based on the views expressed by participants, it appeared that collaboration was stronger between data and analytics teams and areas they serviced, rather than between different analytics teams in an organisation. Interestingly, participants also noted a need to collaborate even wider, with the client base of an organisation. Participants also supported the view that the move toward working more collaboratively did not happen automatically; participants observed that programmes needed to be deliberately put in place to actively encourage this level of change. RP7 also raised the point that data analytics teams needed to have credibility for business to be able to collaborate with them.

5.4.2 Research proposition 2: Resources

The second research proposition as contained in Chapter 3 centres around the resources that are essential in successfully leveraging big data and analytics. This research proposition contains two sub sections which are repeated and analysed separately below.

5.4.2.1 Skills and capabilities

To achieve value out of big data and analytics, organisations need to have the enabling **resources**, which include:

- **technical** skills/capabilities;

Table 5.4.2.1.1: coding for skills

Resources: skills
Skill sets (40)
Overcoming data scientist/IT skill challenge (4)

In coding for skills, the researcher looked for any mention in the transcribed interviews of skills, capabilities, talent or other references to competencies staff needed to possess to be able to assist the organisation in creating value from big data and analytics.

Overview

As illustrated in Table 5.4.2.1.2, the implications on an organisation's technical skills required to leverage big data and analytics were discussed by all participants; with RP2, RP4 and RP6 spending more than 10% of time on this topic.

Table 5.4.2.1.2: Skills and capabilities – themes per participant

Skills	RP2	RP4	RP6	RP9	RP7	RP3	RP8	RP5	RP1	Total
Code count	8	9	4	3	5	3	4	4	3	43
Word count	1519	1132	823	709	490	277	369	317	256	5892
Total	5737	7465	7542	8935	8319	6142	6940	7848	7956	66884
Relative %	26%	15%	10%	8%	6%	5%	5%	4%	3%	3%

The analysis which follows indicated that there was generally a consensus that to create value from big data and analytics, a strong quantitative skillset was required. As an added complexity, the skill sets were not viewed purely as a quantitative ability, but also as business acumen. Participants noted that these skills were in short supply and difficult to retain. It was also acknowledged that these individuals needed to be self-starters who would be continuously upskilled, through courses and training. Diversity in backgrounds and skillsets was also a requirement.

Transcript analysis

According to RP2, the right skills were important in creating value from big data and analytics. Their analytics team were quantitative experts with strong interpersonal skill sets, which were able to engage with business.

> *" They are **a lot more capable of holding their own in terms of the business, or client facing, let's put it that way.** So I find them a bit more business ready, than a while ago."… Data science is a profession that is **drawing people of a different calibre**."*

The absence of a formal, specific "data scientist" qualification and the evolving environment meant that it RP2 aimed to recruit people who are willing to upskill themselves to stay relevant.

*"There is **no formal qualification**, so getting the right people on board [is important]. And knowing what to look for… certain **skills may become superfluous in a short period of time**, so you are **looking for people who have a good quantitative background** and are versatile enough to play in quite a few related fields, and are **self-driven and are going to be acquiring skills through their own endeavours**."*

RP2 also noted that the shortage of individuals with the requisite quantitative skill sets meant that these quantitative resources were in high demand in the financial industry, and priced at a premium. This created a retention challenge.

*"So there is still a **relative shortage of qualified individuals** and you have people who are working in related fields who want to come in, but they don't have some of the basic requirements, but companies are willing to pay up. So, they are **expensive resources** to acquire, and **you are always at risk of losing them** to, not only banks, but potentially Fintechs and the likes."*

RP2 also noted that having a diverse range of skill sets was important. This brought richness of approaches to addressing problems with big data and analytics and contributed to the development of the field.

*"Having people with **different disciplines** within a team brings different viewpoints. I don't think you can ignore that... You can bring in ten people in the same mould, or you can bring an econometrician, with an engineer, with a computer scientist and a statistician, and they **all approach the question differently. You need that diversity in thought** it is still a field that is still very young. So set the tone. **And set the tone is one of inquiry**."*

RP4's comments around data scientist skills echoed that of RP2, however, he reflected on the necessity for a quantitative skill set with strong business acumen.

*"I think for a start there are **simply not enough data analysts**, and I am just going to lump the word data scientist with that. What I mean by that is **not necessarily people who can simply go code, put a piece of code together to analyse data or profile data.** The ability of people to have the **combination having both the business and the technical skill set**. So that's a problem not only faced in South Africa but across the world... I have recently seen that those resources **now have a premium attached to them...**"*

RP6 echoed RP4's sentiments on the importance of business acumen as RP6 viewed quantitative capability as being easier to obtain. RP6 needed quantitative specialists who understood and could communicate business thinking to enable successful collaboration with business.

> *"It requires someone that is, I could use '**well-rounded**' but that sounds terrible. It's someone that **understands business and technical**, because a lot of traditional analytical roles, here's a model or here's your theory, go and build something. And when I try to explain to you about sales or fulfilment process or whatever, you kind of glaze over and you don't really know what it's about. **We have technically-gifted people that struggle to understand business,** and on the flipside, you have business people that struggle to understand technical things. So the challenge, and also the thing that sets us apart, is that we've, I've managed to find the balance for the two. Getting the guys with the **initial technical skill set, but immersing them in business, because at the end of it for me, technicalities can be taught,** 'Here's a textbook, go and read', or 'Here's a class you can attend', but the business stuff you actually have to be a bit more hands on about that."*

RP9 spent time also touching on the challenges around upskilling staff to be able to do a meaningful analysis of data to extract insights.

> *"So, she [data scientist], **understands her business** and she's now obviously **trying to source people.** So again I was taught to analyse data ... **The meaningful data analysis does need training** and help and that is definitely something that we are going to need."*

In contrast, RP7 believed that skills could be obtained from external sources, although at a premium price, rather than needing to be created internally.

> *"I **think we can get resources,** I think they probably at this point in time come **at a premium**."*

On business acumen, RP7 noted that the data analysts required skills that went beyond the quantitative and required a more engaged personality that would also be able to understand the customer's mind set.

> *"We are bringing in a bit of the statisticians to assist, but **they again don't necessarily have the products skills that play that role**… We need to do a lot more to get **actuaries** that sort of understand the products a lot better and be able to close those loops a lot better... Again I think sometimes it's **not your traditional actuary** that you want there, it's **not the number cruncher**… They've got a good understanding of how the world works and be **able to think how our clients think as well**."*

RP3 also agreed with RP6 and RP4, and noted that the ongoing development of quantitative skills was important. RP3's organisation had taken it upon themselves to upskill staff through external training programmes.

> *"Data scientists stem from people with **maths degrees and computer scientists** with **coding** skills. We are trying to **upgrade the skills and develop them ourselves** and we have the first cohort of people that we've sent on this New York Data Scientist Academy... **we are thinking of continuing that**…we've got online licenses for **self-study**, Udemy, Cousera; things like that."*

RP3 also identified skill shortages and retention challenges as key.

> *"I think because t**here's scarcity of skills, retention of the skills, once you've acquired it is also a challenge**, because our data scientists are well sought after, after we upskill them, we need to **make sure we pay them right** and keep them happy, keep them interested in whatever interests them… **There's so much work and there are limited resources…** we have **vacancies** based on the book of work we've got; we need more staff and we've got budget for it but **we can't find resources**."*

RP8 also noted a shortage of skills and suggested that introducing diversity into teams was a manner in which the current shortcomings may be overcome.

> *"…[analytics] **people don't necessarily always understand the business very well**… I think if we get **different people, mix the old and the new…** I think getting different people, they could be people from aviation, they are going to ask questions about our industry that we will not ordinarily ask because we take things for granted and I think having that **diversity of experience, thought, people that would be great**. Whether its gender, race, sex, professional skills and then obviously find ways of interacting with the market."*

RP5 agreed that diversity is a key factor but, like RP2 centred around quantitative skills, which remain the bedrock of skill sets required:

> *"Probably to get this full sets of skills you employ a diversity of people with **degrees in maths and statistics even engineers** as I'm an example, people who studied physics etc. that can do statistics."*

RP1 also believed quantitative skills are a must in the big data and analytics journey.

> "***Quantitative skills are key, data and quantitative** skills yeah. And it's a mind-set."*

When queried on the need for business acumen skills, RP1 noted data scientists would not need that as a business person would be present to assist.

> *"It's extremely helpful but **it's not a necessity,** because you've got someone saying here's the business problem to solve."*

Conclusion

The findings on skill sets are captured in a Table 5.4.2.1.3.

Table 5.4.2.1.3: Interview findings related to skill sets

Resources: Skills	RP2	RP4	RP6	RP9	RP7	RP3	RP8	RP5	RP1	%
Quantitative skills (mathematics, computer science, statistics, coding, actuarial)	X	X			X	X		X	X	67%
Business acumen/understanding	X	X	X	X	X		X		-X	67%
Ongoing upskilling	X			X		X				33%
Diversity	X						X	X		33%
Skills shortage: Difficult to attract/retain	X	X			-X	X				33%
Premium remuneration	X	X			X	X				44%

Overall, respondents noted that different skill sets were required to enable an organisation to create value from big data and analytics. The majority (six) of the participants noted that it was important that data scientists possessed quantitative skill sets. In addition, six participants made specific reference to a need to combine quantitative ability with business acumen to succeed in generating value from big data and analytics. However, one participant (RP1) did not believe that acumen was required as the business practitioner the data scientist collaborated with would possess this ability. The need for data scientists to be self-driven and upskill themselves were both noted by three participants, as was the need for diversity in skill sets.

In addition participants also noted other complexities. Three participants noted that recruitment and/or retention of the relevant skill sets was challenging. The same participants also mentioned that there was a need to pay these employees premium salaries. One participant (RP7) noted that even though skills shortage was not a challenge, data scientists demanded a premium.

5.4.2.2 Financial resources and physical infrastructure

To achieve value out of big data and analytics, organisations need to have the enabling **resources**, which include:

- **Financial resources** and **infrastructure.**

Coding for financial and physical infrastructure

Table 5.4.2.2.1: Coding for financial and physical infrastructure

Resources: other
Funding (8)
Infrastructure (24)
Overcome infrastructure challenge (1)
Resource: data warehouse (4)

In coding for financial and physical infrastructure, the researcher looked for all references to the different resource participants mentioned they need to acquire to be able to pursuer the big data and analytics journey. The researcher also looked for reference to any other resources that would be required.

Overview

From analysis of the transcripts, it became apparent that in addition to human capital resources changes, the participants also noted that their big data and analytics journey introduced IT and financing demands. Resources did not garner a great amount of discussion, as indicated in Table 5.4.2.2.2.

Table 5.4.2.2.2: Other resources- themes per participant.

Resources – other	**RP9**	**RP5**	**RP2**	**RP3**	**RP8**	**RP6**	**RP7**	**RP4**	**RP1**	**Total**

Quote count	9	8	3	4	4	2	3	1	1	35
Word count	1184	985	309	337	343	343	335	50	49	935
Total word count	8935	7848	5737	6142	6940	7542	8319	7465	7956	66884
Relative %	13%	13%	5%	5%	5%	5%	4%	1%	1%	6%

Transcript analysis

Participants dedicated more time to the topic of the physical infrastructure required to be able to leverage big data and analytics, and is therefore discussed first.

IT infrastructure

RP2 noted that advances in processing ability had facilitated the big data and analytics journey.

> *"A lot of what is now possible for us to do, or a lot of what is being done now, is being facilitated by the fact that* ***we have access to creative processing capability, better software, more storage and this domain of statistical analysis has become accessible to people****."*

However, RP2 noted that organisational policies hindered the ability of the organisation to keep up with the latest technology trends and to obtain the appropriate infrastructure operate effectively to create value from big data and analytics.

> *"So, I am talking about things like having* ***access to the right software, memory, storage capacity, the speed of access****.... **We find that, we are often curtailed by things like software standards for the enterprise** that require that you use certain software stacks, certain programs, and they don't work for us."*

RP3 noted that the integration of disparate databases was problematic and hence the infrastructure related to this was a focus area for their organisation.

> *"So whilst you're trying to build the tools, we're trying to build a* ***platforms in which we can land all our data,*** *being a bank with lots of history and tech debt, and data debt… all of that is the main focus now* ***where we build all the infrastructure*** *off which the future hubs can live."*

RP5 emphasised the need for updated systems and infrastructure to replace the legacy main frame systems to ensure that the big data can be processed efficiently and in real-time rather than batch processing.

> *"Some of the* ***legacy issues*** *we have in the bank that is still constraining us, is for example* ***core transactional systems*** *are still Cobalt-based mainframe. A lot of the newer, more advanced models and more refined data approaches we house in different data environment... These are typically the* ***platforms that have got a lot of data and distributed computer resources*** *… I think t**he IT investments on the data platforms**.. are* ***big components in terms of the journey.*** *We don't have modern database architecture that allows main frame streaming, but I think we can overcome it."*

RP7 concurred with the concerns RP5 expressed around the legacy infrastructure, and noted that their organisation was focussed around consolidating systems.

> *"In my world specifically, a lot of our time spent on system consolidation… So* ***you can get access to the software quite easily but to start using it I think is not that easy****."*

RP6 also noted that the size of the organisation and the time it took to transform presented a challenge to the ability to getting right infrastructure in place to support big data and analytics.

> *"So you* ***need to also try and figure out how to get data centrally located that talks to infrastructure.*** *What should the infrastructure be?… the challenge comes in, because we're such a large organisation .. So because* ***you don't have the speed for getting there from such a large organisation****."*

RP8 went one step further, and noted that the old technology had a cost implication as it was a hindrance to becoming a fully digital organisation.

> *"Now let me put it to you this way…* ***we've got old technologies****... If we had to start from scratch and we had a magic wand, we could probably do the same work for half the cost and maybe half the people because of data, digital presence and the* ***ability to get data and information to anyone anywhere in the world anytime, in the universe for that matter****."*

RP9 cited duplicate systems and the inability to easily reconcile the data in these as a challenge. Similar to RP5 and RP7, her organisation was seeking to overcome this legacy challenge.

> *"So, **we've got multiple systems, even within a business unit, we're trying to consolidate**…So, I think the way you build that is so absolutely critical and as I say **integration is absolutely key.** You can't be manually managing these interfaces anymore."*

RP1 noted that the data landscape was different in that traditional data warehouses were being replaced by more non-traditional data storages such as data lakes.

> *"So although **IT needs to understand architecturally how they are going to design** what their **data landscape** looks like in the traditional or non-traditional way. So traditional would be to build a warehouse and **non-traditional** being is the more **agile data warehouse with my data lakes and my big data** strategy."*

Financial resources

In addition to physical infrastructure, participants also noted financial challenges. RP1 explained that cheaper, cloud-based storage created an expectation that the data and analytics journey would not have incremental financial demands. However, the new infrastructure also came at a cost, and this cost needed to be recouped.

> *"As much as people say, yes you've got the Hadoop and it uses cheap storage... It still uses storage, **it still sits in a data centre, it still consumes power**. So if **you're just producing nice free dashboards of meaningless information that no one is monetising... there's no point**."*

RP3 agreed with RP1, and noted that a return on investment was needed, whereby stakeholders first had to demonstrate the value that big data analytics projects could deliver, before organisations should be willing to invest.

> *"Initially the problem was when we try to prove the possible value add of the embarking on this plan and strategy for the bank**, there was scepticism because we required some money** and some funding and it was '**What will be the return of value**? We can't really see the benefit add?' And we had to go away and build some models and tools*

*that can **try and demonstrate the possible new value add** that the bank wouldn't otherwise realise."*

In contrast, RP8 noted that the exploratory nature of some of the big data analytics projects meant that a new mind-set was required and that they had to come to terms with the investment required, despite the difficulty in demonstrating value upfront.

*"And it's like going out there to explore for minerals or treasure out there in the sea, you either find it or not find it, it costs a lot… But also I think just **leadership** saying; 'Okay, **we are going to have the courage to spend money and not necessarily expect immediate return**'. Mining for the insights is just one of those things."*

RP9 was aligned with RP1 and RP3 and noted that the organisation always had financial challenges and that a demonstration of revenue, or reduced costs was required to justify an investment in infrastructure.

*"But we've always got money and people challenges… from a leadership perspective you know we have to get everything approved... So, **no matter how sexy or exciting** or how much you think you need it, if it's not going to get you **revenue uplift or significant efficiencies, they won't fund it, they won't allow you to do it**. So, it's a challenge".*

Conclusion

Table 5.4.2.2.3: Interview findings related to skill sets

Resources: Other	RP2	RP3	RP5	RP7	RP6	RP8	RP9	RP1	RP4	%
IT infrastructure	X	X	X	X	X	X	X	X		89%
Financial requirements		X				X	X	X		44%

Based on the above, eight participants noted an IT infrastructure implication. Participants noted that financial institutions had legacy issues and that the outdated infrastructure caused a barrier to the successful implementation of big data and analytics programs. Organisations required different infrastructure that could consolidate disparate data and systems. In addition, the point was made that implementing the appropriate infrastructure could take a significant amount of time due to the size of an organisation, while outdated organisational policies did not permit the implementation of technology stacks that were required.

In terms of financing, four participants noted that the big data and analytics journey required financing – this went beyond simply the infrastructure costs, although this was a big component noted. Three participants noted that traditional return on investment considerations applied when embarking on the big data and analytics journey and that the value needed to be demonstrated upfront to secure funding. One participant held a different view - that investing in big data and analytics capabilities required a different mind-set in which organisations simply invested in the hope of generating some form of return, which may not be known or quantifiable in practice.

Overall, the findings support the fact that it infrastructure and financing resources are important additional requirements for successful creation of value from big data and analytics.

5.4.3 Research proposition 3: Value

The third research proposition as contained in Chapter 3 is repeated below.

> To create value with big data and analytics, organisations need to be able to **articulate** what **value** it is looking for, as this enables articulation and tracking of success **metrics**.

Coding for value

The coding process identified the codes in Table 5.4.3.1 as relevant when discussing value, based on the various definitions attributed to it by participants.

Table 5.4.3.1: Codes related to value.

Value		
Analytics sale (1)	Empowering customers with analytics (3)	Improved customer experience (16)
Collaborating with clients (1)	entrench customer (2)	Monetising Data (8)
Competitive advantage (12)	Facts, insights for decision making is in and of itself value (2)	Single view of customer (7)
Customer lifetime value (1)	Financial impact (5)	
Data as an asset (2)	Fintech competition (6)	

In coding for value, the researcher looked for any reference made to outputs or benefits the organisation aimed to achieve as a result of using big data and analytics. This included any references made to the metrics used in measuring these outputs or benefits.

Overview

Analysis of the transcripts revealed that the concept of value and mention of value attracted a fair amount of attention in discussions, as indicated by the distribution in Table 5.4.3.2.

Table 5.4.3.2: Value – themes per participant

Value	RP8	RP1	RP3	RP2	RP9	RP5	RP6	RP4	RP7	Total
Quote count	11	11	7	4	7	6	1	3	4	54
Word count	1708	1127	748	535	726	413	283	244	288	6072
Total words	6940	7956	6142	5737	8935	7848	7542	7465	8319	66884
Relative %	25%	14%	12%	9%	8%	5%	4%	3%	3%	9%

The general view expressed was that the value to be generated by big data and analytics would be in the form of financial impacts, improved customer experiences, improved competitive advantage, operational efficiencies as well as the potential sale of data and/or analytics as goods or services.

Transcript analysis

In articulating value, RP8 spoke optimistically and futuristically about the use of data and the new revenue streams it could create.

> *"What we can get out of data and **when we have a history** of being able to get value out of data and are able to **monetise data**, and therefore **generating new forms of revenue** for our business, I think then we have hit the spot."*

RP8 envisioned the value big data and analytics could create as a better customer understanding.

> *"Question is; **what is the next big thing**? And I think that **sits in the data,** somewhere in the data… how do we get to the psyche of the customer collective in terms of what is the one thing that if we did, it would be the **Uberisation of our market?**"*

RP1 stressed competitive advantage as a concept of value and noted that organisations needed to explore and exploit their big data, or risk other organisations, notably Fintechs, doing so before they do.

> *"If you don't do this someone else will… We are a data company, so if* ***you are a data company*** *you either* ***understand your data*** *very quickly or there will be some* ***Fintech start-up that will understand it better than you, sooner than you and then will undermine your ability to compete****."*

Despite this, RP1 also emphasised that organisations should not aimlessly endeavour to create new insights through data and analytics, but needed to know what the intention was with creating insights and how it would be able to generate financial benefits for the organisation.

> *"If you have the ability to get a lot of data and to* ***make sense of that data before your competition*** *then you're able to* ***make better decisions*** *or* ***you're able to monetise it more effectively****…"* ***"How much more money are you going to make now that you know this?****... So if you're just producing nice free* ***dashboards*** *of meaningless information that no one is monetising...****there's no point****."*

RP3 found that being able to understand and anticipate customer preferences would generate revenue as a source of value.

> *"We spend analytics by analysing each of our* ***customer's behaviour and their preferences****… you know we can* ***recommend products*** *to them at the right time... Then* ***customers feel like the bank cares,*** *'They know what I like'. So that indirectly comes through to us as more client* ***retention*** *and* ***attraction of new customers*** *were to come through we would have new* ***growth****."*

RP3 further noted that data was an asset that had value, but it was unclear what that value was and how it could be realised.

> *"People know that there is value, definitely has to be –* ***data has to be an asset*** *– but* ***there's still no clarity*** *on, you know convincing* ***that we have realised the value yet****."*

As a step toward realising some form of value, RP3 discussed how their organisation was making their analytics tools available to existing business customers.

> *"So besides the offering this free, we have embarked on a decision of, **how do we realise value of different types of tools and things we're doing**?… Some **tools we may want to sell to our clients**… When we combine data [with our customers] and we give **increased insights to them [customers]** for their customers, they can use it to improve their market share which indirectly will come through to us because **they bank more with us**, which makes more money for us. **It's a win-win**."*

RP2 also saw value as a financial impact although he conceded that strategic motives such as pricing to be more competitive would also form part of a decision and hence competitive posturing was also seen as value.

> *"I see value on a few axes. I think the natural one is always a **financial impact**. So there is a business problem to be addressed, and you do some analysis which proves or disproves a hypotheses or helps generate a recommendation **and the business changes as a result,** and there is a **financial benefit. But, financial benefit cannot be looked at in isolation**…. For instance, we may want to change your pricing structure because it is uncompetitive, as an example**. We feel that the long term benefit outweighs the short term revenue sacrifice,** and to remain kind of sustainable as a going concern. So you can't look at that in isolation."*

RP9 conceived of value as the achievement of greater operational efficiencies.

> *"So, again that's our challenge is understanding where as a broker we need to use technology **more efficient** but not to lose our role as the broker."*

RP9 noted that their organisation would not be able to or wish to sell data to insurance companies as they would face resistance in the market.

> *"…an initiative to **sell data** back to insurers and this is also one of the contentious things, like w**hy should they be paying for it?...** I don't see us selling data locally at any point in time."*

RP5's commentary around creating value centred on better segmentation and more delivery of goods or services to the customer experience and improvement of it.

> "*My third point was around the **customer-centric delivery platforms** with the customer segment of one, which is very powerful. And we think that the more progressive approach taken by Bank W… puts us **in a good position to be a beneficiary of developments** in the new… let's say the world of data and analytics… we are now in a position where we are almost ready to roll out a new organisational model which starts to **bring in much more front-end client experience kind of stuff**.*"

RP6 said that the value his organisation strove for was increased cross selling (a financial impact) and entrenchment of customers

> *"Effectively as a team we have, let's call it two main objectives… So how do we increase cross-sale ratio, and correspondingly, how do we **increase life-time value** of our customer base?"*

RP4 referred to the value of big data and analytics as an enhanced understanding of customers through the insights generated.

> *"'That customer approach can only be done if I truly understand who my customer is and if **I'm truly using the powerful insights generated both in descriptive and predictive [analytics].**"*

RP7 stated that the value in data and analytics lay in obtaining an understanding of desired client behaviours that would generate more revenue and growth.

> *"So for us I think the challenge is it's going to be for quite a while: how do we understand the data a little bit better and actually **model the behaviour you want from clients and actually reward clients appropriately for the right behaviour** or get them to actually want to **buy more products** from us because we are actually providing them with a **good service**…."*

As it relates to specific metrics used in measuring value, the following two comments were made.

RP7 noted that the lack of understanding of how value was created presented a challenge to demonstrating value.

> *"I just think that **our biggest challenge is to really be able to sort of understand where we've actually made successes.** …So being able to really **understand where our actions and how our actions actually translated into a particular client and shareholder outcome… I think it's still a challenge.**"*

RP2 supported the view expressed by RP7 and noted that his organisation had encountered a problem where a lack of definition of a metric and volume of success meant that the success of the tools could not be proven.

> *"So that we can say 'Well now that you have started using our tool, your sales or your **productivity has increased by X% purely due to our value-add',** so how do you measure that? And they'll say '**No. We were going to grow anyway**'. So, **how do you split out the normal growth from the extra growth because of our tool?** We want to show the return, the value as well."*

Conclusion

Table 5.4.3.3 provides a summary of the above analysis.

Table 5.4.3.3: Summary of participants view on value

Value	RP8	RP1	RP3	RP2	RP9	RP5	RP6	RP4	RP7	%
Financial impact	X	X	X	X			X		X	67%
Competitive advantage	X	X				X				33%
Improved customer experience	X		X			X	X	X	X	67%
Operational efficiencies					X					11%
Data and analytics as goods or services		X	X		-X					22%

It was clear that all participants were able to articulate what their organisations viewed as value, and therefore had a sense of what they are pursuing through their big data and analytics journey. The four broad categorisations of value, in order of support, were:

1. Financial impact tied with improved customer experiences.
2. Competitive advantage.
3. Data and analytics as a good or service. As represented by the –X in Table 5.4.3.2, RP9 disagreed with this concept and raised the point that the organisation expressly did not see that it could sell big data to its clients.

4. Operational efficiencies.

However as RP7 and RP2 stated, organisations had an inability to measure and demonstrate precisely how and what value was realised. The researcher noted that none of the research participants articulated metrics that their organisations applied in measuring value.

5.5 Other findings: Maturity in big data and analytics

At this juncture, other issues that were raised by numerous participants will be discussed. These issues were not part of the initial scope of this research, but provided some insight into the research findings. The researcher noted the gravity with which this was addressed and therefore, although this overlapped with decision making, considered that it warranted separate discussion. These findings are presented here, and are discussed more fully in Chapter 6.5.

Getting the basics right: data

In many instances, it became apparent that the organisations included in this research made use of their big data and analytics to do reporting. The analytics applied was therefore descriptive, rather than predictive analytics. This was verbalised in various ways by all participants apart from RP5, who was the only participant that believed their organisation used data and analytics extensively. However, the researcher observed that RP5 often cited credit decisioning as a reference point in discussing the organisation's use of big data and analytics, indicating that this organisation may well also be using data at a more elementary level.

RP1 offered a definition of the difference between business intelligence (BI), which he saw as synonymous with management information (MI). RP1 noted that reporting and doing analytics on historic information was not the same as running predictive analytics and that he believed that most of what his group was doing was analytics on historical information.

> *"The one thing that you need to be clear in your mind is that* ***there is a very big difference between BI/MI*** *which are basically one of the same things. "Business intelligence/management information"* ***versus an analytic****;* ***most of what Group is currently doing sits under business intelligence.*** *If you can simply go and do… 'if, then else', on historical information then you are doing BI and MI, you're not doing [predictive] analytics."*

RP1 explained that an organisation needed to be comfortable with their data before they could embark on a predictive analytics journey. His organisation was on a journey to consolidate their data and, as an organisation, to obtain a more complete understanding of their clients.

> *"Because you're looking and saying 'If the client is the centre' ...**we need to make sure we understand more about that client... so that's MI and BI…** unless you're happy that all your **client data** is accurate then you can't really do any analytics."*

RP3 observed a similar trend of doing descriptive analytics in his organisation, where 80% of his team's time was spent on business as usual (BAU), or what RP1 referred to as BI. The remaining 20% of their time was spent on more novel activities, which was where predictive analytics would play.

> *"We try to ensure it's an 80/20 split, where **80% is BAU stuff and 20% is trying to work on the pie in the sky kind of stuff**… The more proactive stuff that we can showcase."*

RP2 and RP9 noted that a split of functions between teams aided in ensuing that the predictive analytics agenda was pursued.

RP6 did not have the same split teams or clear dedicated time allocations to various objectives and, as a result, reported a low rate of producing anything other than BI.

> *"Because **we focus a lot of business as usual things**… it becomes very difficult to try and implement anything new… **The new run rate has been very sort of mediocre** or whatever five out of ten type of things that actually go from implementation going forward."*

RP8 noted a similar problem in his organisation, where "urgent stuff" related to projects that currently make money, took precedence over the "important" act of looking for "gems" in their data.

RP1 speculated that organisations used hyped-up terminology to describe their activities. This implied they are undertaking predictive analytics, when in fact it was descriptive in nature. The unintended consequence was that these organisations appeared to be more mature than they actually were.

> *"**If you don't use the word "analytics" in something, then people aren't going to buy your product**, if you just say it's BI and MI then people are going to know that, that's old school... you say "big data" then they get excited. You have to use the buzz words…"*

RP4 offered a slightly different perspective and noted that it is not just "window dressing" as RP1's observation implied. Instead, he believed the big data and analytics move had genuinely improved management information.

> *"So it has moved away from traditional monthly reporting that I produce to 'What are the real insights within that reporting?'… It is about making it to the next level so, '**How can I bring disparate sources of data together even within an MI** (management information) construct and how **can I start showing insights out of that data between seemingly unconnected data points?'"***

The points raised by participants illustrated the fact that data and analytics were used largely in a descriptive manner, to enhance management reporting and insights. It is not yet primarily used in a more evolved manner. A reason for this was challenges with data and the need to getting the basics right first before more advanced analytics could be pursued. Getting the data in order was a key part of getting the basics right first.

A point of discussion raised by participants was that they needed to have a level of faith in their data, before the benefits of big data and analytics could be pursued to create value. Part of their objectives was to consolidate disparate data sources and to ensure that data quality was established.

RP1 noted that having the right data was a cornerstone to being able to produce trusted analytics.

> *"… The basics are not so easy to get right… **If you actually don't have faith in your data:** the quality of it, the source of it, the velocity of it... **how do you trust the analytic that comes out**?"*

RP4 agreed with RP1, but offered sage reflection around the need to balance how the data issues were addressed to ensure that the organisation did not compromise its upkeep with competitive forces due to a pre-occupation with fixing its data. Instead, it needed to figure out a balance between fixing the data while pursuing a big data and analytics initiative as a source of value.

> *"And I don't think that you necessarily have to have your data world sorted out before you can hit high levels of maturity because my feeling is that you will then just be going in a constant circle…* ***In colloquial terms, how I do fix the plane while I'm flying it****? So how do you recognise that there are data touch points that are going to require 100% data quality and those that will require a lesser level of quality?"*

RP5 picked up on the same point and noted that one initiative undertaken in his organisation that contributed to their success in using big data extensively, was that they focused on consolidating a substantial amount of, but not all, customer data across the bank on a central database.

> *"We put together big data in the bank around the customer-centric approach… customer…* ***We have about 99 and a half percent coverage I think, of individuals … on juristic entities the number is I think about 96%. They are still relatively high****. That's one of the* ***building blocks and what that have allowed us to do customer-centric*** *product process...."*

RP3 expanded on the legacy issues of disparate data sets that did not conform, as it contributed to the challenge of getting all data in the same place. Trying to "land the data" was a time consuming process which caused delays.

> *"So whilst you're trying to build the tools, we're* ***trying to integrates all 300 systems that we got into one central Hadoop stack and conform the data****,* ***integrate data, clean the data*** *…* ***We're trying to land data****. So without building the infrastructure and the rails, you know you can't take advantage of that* ***because it's time taking****."*

RP7 echoed RP5'and RP3s sentiments on the importance of a consolidated view of data on customers, but also emphasised that in addition to bringing together disparate data, data quality was also an issue.

> *"I think one of the other challenges we have in the insurance world,* ***especially if you've been around for a few years as well, is quality of data****. Even something that seems simple like a single view of the client across the group… becomes a struggle."*

RP9 built on this and noted similar issues with disparities between data in their data warehouse and the source systems having caused a lack of trust in the data in the system, and noted that the only way to restore trust was through fixing the data.

> *"For years had this core data warehouse we only have a bit of information in it, and then we've got a whole lot of other source systems sitting under it and that's c**aused utter chaos because you've got this constant mismatch between data**… The power is going to be in the data at the end of the day, **if the data is still a mess, then we are still going to have trust issues…** success breeds success, so if we can start getting the data right I think the trust will follow."*

Based on the above observations by participants, organisations were focussed on remediating data issues, many of which was a legacy of not having viewed data as an asset to be curated. This preoccupation with data quality and data consolidation was necessary to ensure that trusted analytics could be produced, so that analytics could evolve in tandem with business' need for deeper insights.

The researcher expected participants to refer to the use of unstructured and external data usage as part of their decision making process. However, this was not the case. Participants generally noted that due to the legacy challenges with data, they were not yet far enough advanced in their big data and analytics journey to make meaningful use of substantial amounts of unstructured or external data.

Three participants in this research (RP3, RP6, and RP5) noted that their organisations made use of unstructured data, while RP4 noted he had observed this trend. It was surprising that RP3 and RP6's organisations used unstructured data given that, unlike RP5, their originations did not use data extensively. However, these surprising factors may seem more expected if one considers that a "building the wings while we fly" mentality has taken root. Organisations were eager to gain traction and prove the strength of big data and analytics so as to garner more broad-based buy-in and adoption.

RP5 was the only participant who noted extensive use of big data and analytics in decision making, and strong leadership buy-in. In terms of the use of unstructured data, he noted that his bank typically uses structured data, although it had unstructured data as well which it is starting to analyse.

> *"We are typically **working on structured data a lot more than unstructured**. Other unstructured data we have for example is voice recordings… We are voice transcribing on an automated basis the bulk of all voice recordings. In service call centres we have transcriptions of calls. Now, that is **a bit more tricky** … **natural language processing on that to extract meaning or the like from that is something that we are not very mature at; but we have embarked on it**."*

RP5 noted that his organisation used external data to augment internal data in making credit decisions.

> *"If you don't integrate in to bureaus and use external data, you wouldn't be able to take decisions with customers when you don't have a past history."*

RP3's organisation appeared somewhat more advanced and has built, and will launch, a credit scoring model using unstructured data. They were also using external sources to augment their own internal sources.

> *"So firstly, **unstructured data, that we are looking at** such. To our social media data, we're trying to use all of those things to re-inform and challenge some of the traditional ways of thinking such as alternative credit models and we set data to give people credit, but newcomers who haven't had debt before are turned away because there's no history on them… Their social media data, other data… We can still give you credit without you having history, **that's something new that's being launched** because **we also have data partnership agreements with external credit providers**, **structured data that we are buying data to enrich our sources, sources of data, maybe bureau data**."*

RP4 noted that he had started to observe the practice of using unstructured data as well as sourcing external data to augment internal data, but that the use of unstructured data was not at a very advanced level.

> *"And what we are starting to see now is not just non-traditional internal data sources being used to inform outcomes but also external sources being used to inform outcomes and in potentially seemingly unconnected data sources being used. And I talk there about both structured and unstructured side… What we are certainly starting to see is the use of the unstructured data, using machine learning capability, but I think it is still at a very low level of maturity."*

RP2 noted that his bank uses mainly structured data.

> *"No, predominantly structured data. In fact, exclusively structured data. Even in the most progressive of our three streams, that will be the insights team,* ***there isn't much unstructured data that comes into play.*** *I can't think of any off hand. I mean, we have plans to but, to be frank,* ***there is a lot of unstructured data that we don't tap into****."*

RP2 further explained that there are plans to start using unstructured data once the structured data is exploited more fully.

> *"But there are many* ***opportunities to augment that data*** *with other information that is available from other product houses or let's call it things like statutory returns; there are* ***things like cross-border payment monitoring*** *that is done in the Group corporate centre. And none of these data have yet been appended to our existing data sets or collated or injected into our data landscape. So* ***that would be our first hurdle, and we have plans to do that before we go into the unstructured realm****."*

RP8 also noted that in his organisation, data was mainly structured and was used mainly for risk management and was not fully exploited.

> *"I think this business being mainly about managing risks that's all about data that's mostly about structured data and you can still build a lot of insights from the data that we have."*

RP9 similarly shared a view that only structured data was used "because it is bad enough just with the structured data", indicating that her organisation was still coming to terms with structured data.

Similarly RP7, RP6, RP9, and RP8 noted that they enriched their internal data with external data, mainly of a structured nature. RP8 also voiced aspirations to start incorporating social media, which represents unstructured data.

Based on the above, the majority of the participants spoke of an ambition to progress to using unstructured data in due course, once they had gotten the basics right and were fully able to exploit their structured data. Similarly, the majority of participants used external data to augment their internal sources, but it was not used as extensively as internal data.

Qualitative methodologies

The researcher had expected to find that organisations also made use of qualitative methodologies in their decision making, however, this turned out not to be the case. Only one organisation, RP6's team, made used qualitative methods to gain insights.

In terms of using qualitative methodologies, RP6 noted that one of the teams he led undertook research by engaging directly with consumers to gain an understanding of their perception of the company's products. RP6 was the only participant to refer to the use of qualitative information obtained directly from consumers to be used to inform decisions in conjunction with the quantitate data.

> *"We are, to my knowledge, the only team in the entire bank actually, that has a research team incorporated within analytics. So analytics has very much been about quantitative data… Research will actually* ***go out there and survey customers****, in order to figure out what is our comparison to other banks, in this example. So, how do we do company analysis? Figure out which products are suitable to our customers… So incorporating research into the analytics function has enabled us to look effectively at both sides of the coin.* ***It's not big data stuff,*** *it's as opposed to* ***quantitative and qualitative, and how do we combine those two together quite effectively?*** *That actually gives the true value of insight as opposed to just saying:* ***'Here's what the numbers told me and then go forth', and then you actually shoot yourself in the foot.*** *That's what we do as a team…"*

RP6 also noted that publicly available information such as that on social media was used at a basic level, such as complaints monitoring, but the organisation was yet to escalate beyond this.

> *"So* ***social media analysis people have done normal stuff*** *in terms of filing complaints and finding a service issues, right. So* ***that's the basics so we do that…****"*

Based on the above, RP6s organisation was in the initial phases of exploring the concept of thick data and qualitative research methodologies. They were gaining a richer understanding of the insights they could achieve by appropriately combining thick data and qualitative methodologies with big data and analytics. However, no other organisation voiced a similar level of evolution in their approach to gaining new insights, particularly around customers.

Adoption challenge

A further point raised by participants related to the challenges they experienced in getting business to adopt big data and analytics. A lack of adoption was evident in the finding that only eight of the nine participants in this study noted that data was used extensively in decision making within their organisations. Change management also surfaced as an issue which hindered collaboration, as discussed Chapter 5.4.1.3.

RP2 explained that business resisted using tools designed by big data and analytics teams as it was perceived as a criticism of their work. He offered a view that the data and analytics team should include stakeholders early in the journey to make them feel like they are partners, rather than adversaries, in gaining better insights.

> "*However, some of the other tools there's still the problem of change management which is the biggest issue;* ***of people willing to use stuff that you actually work with****… The reason is,* ***sometimes the sort of stuff you show them, indirectly comes across as 'the stuff that you have been done all along is wrong, its sub-optimal*** *and here is a new and cooler way of doing things', and these are of new insights; and the people that are listening obviously****, become defensive.*** *Obviously, the way you deliver the message,* ***try to involve them in the process early;*** *let them realise this is a partnership approach."*

RP9 agreed with RP2 in that their organisation was looking to implement data and analytics more successfully by including the stakeholders in the journey.

> *"I think what also happened in* ***the early days is the stakeholders weren't involved****… Our approach now is way* ***more inclusive, way more about stakeholder management****."*

Offering a further suggestion to improve adoption, RP4 noted that it was important to create a burning platform by, for example, using a strategic objective as a driver of change.

> *"In our view y****ou need to use the low hanging fruit*** *to drive the imperative … agenda. So I don't think there's one-size-fits-all but definitely finding a balance between the moving parts and* ***understanding the organisations strategic imperative*** *and using that imperative to drive the balance, is what we are seeing."*

RP5 noted that having a business case aided his organisation in becoming more data-led.

> *"In many of the business unit's **credit was the lever for unlocking analytical capabilities** because there was a **huge financial case** around doing it and people were willing to hire relatively senior quantitative people into those roles which then facilitated growing this into other areas."*

Five participants raised the issue of legislation or regulation in South Africa, which was compelling financial institutions to pursue a big data and analytics strategy. This included legislation, such as the Protection of Personal Information Act (POPI), banking regulation, such as the Basel Committee on Banking Supervision's regulation BCBS 239 and accounting regulation such as International Financial Reporting Standard (IFRS) 19 on Impairment of Financial Assets such as loans, and IFRS 17 on insurance. The role of legislation or regulation in promoting an agenda for creating value from big data and analytics was not part of the original scope of the research but emerged as relevant during the course of this research.

RP4 noted that organisations used regulation as levers to not only achieve compliance, but to go beyond that and build a sound big data and analytics capability.

> *"What has become very relevant now is the regulatory side. So, **how do I ensure that I get best bang for my buck when I go down the regulatory side** and solve the regulatory challenge?…There are some banks that are **looking to the regulatory driver to drive their end to end enterprise data capability** because they know that the regulation: they don't have a choice but to comply with the regulation but let's not just use this as a tick box exercise but let's start to achieve what we coin the term as '**the regulatory dividend**'."*

RP5 noted that regulation was used as driving force in his organisation.

> *"So breaking down these barriers deliberately and getting people together in forums, that has been a very big drive. With BCBS239 we realised that we could not allow the separation of data between finance and risk, we couldn't meet the principles. That **drove a nice set of alignment…**"*

RP8 agreed and noted that the regulatory changes would persist.

> *"We've got some **legislative changes** in the industry coming, which are going to **force the industry to start using data better**."*

The opposite side of this “regulatory dividend” was explained by RP7, who noted that a variety of factors impact the prioritisation of big data and analytics and legislative changes demand a significant amount of time, diverting resources from this journey as opposed to supporting it.

> *“There’s lots of driving force that drive our decision from a prioritisation perspective. We actually live in quite a complex world and at least a third of my implementation team’s time is spent on legislative changes.”*

Participants noted that an imperative for change was needed as this helped create focus and broad-based buy into big data and analytics, across the organisation.

The implications of this, as well as the other findings that were presented in Chapter 5, is discussed in more detail in Chapter 6.

Chapter 6: Discussion of results

6.1 Introduction

The analysis undertaken in Chapter 5 enabled the researcher to establish evidence for, or against, each of the research propositions that were set out in Chapter 3. These findings are presented in the various conclusion sections contained in Chapter 5.

In this chapter, the research findings set out in Chapter 5 will be related to the literature reviewed in Chapter 2. Therefore, this chapter aims to make sense of the findings resulting from the analysis of the nine in-depth interviews that were conducted with the leaders of financial services organisations with the information gleaned from academic literature.

For ease of reference, each subsection commences with a presentation of the individual research propositions from Chapter 3. Thereafter, a summary of key literature from Chapter 2 is presented, and linked to the findings in Chapter 5 to establish if there is agreement or dissonance between the research findings and the literature. Finally, this is all pulled together to assess the implications for the conceptual model of the process by which big data and analytics is used to create value as presented in Chapter 3.

6.2 Discussion of research proposition 1: Decision making

6.2.1 Leadership support for data-led decision making

The relevant extract from research proposition 1 is reproduced below. It relates to the expectation that big data and analytics will result in data-led decision making.

To create value from big data and analytics:

- Organisations need **leadership** that **support** the use of big data and analytics in decision making;

In light of the view expressed by DalleMulle and Davenport (2017) that financial services used big data strategically, the researcher had expected the financial institutions represented in this study to be well advanced in their big data and analytics journeys. However, based on the research findings contained Chapter 5.4.1.1, although all nine financial institutions represented in this research used big data and analytics to enable data-led decision making, only one organisation used it extensively. The remaining eight financial institutions had data and analytic capabilities but participants acknowledged that it was not used extensively, and offered potential reasons for why this was the case. The finding that big data and analytics was not used extensively was incongruent with the expectations formed by the researcher based on the literature reviewed by DalleMulle and Davenport (2017). However it was consistent with some of the other literature reviewed. McAfee & Brynjolfsson (2012) noted that convincing executives who rely heavily on intuition to become more data-driven was a challenge. Ross et al. (2013) studied more than 51 organisations and found that there are few companies that use data consistently to make decisions, despite the fact that companies that did use data-enabled, evidence-based decision making were more profitable than their counterparts. They noted that to become data-led, the organisation needed to undergo a disruptive culture change to concisely apply data-led, evidence-based decision making. This was supported by Vidgen et al. (2017) who noted that a point of departure in leveraging big data and analytics was creating a "big data and analytics strategy"'. RP4 and RP6 noted that clear big data and analytics strategy was absent in their organisations, thus leaving the organisation unclear as to how and for what purpose big data and analytics should be used.

Ross et al. (2013) further postulated that only once decision makers had data at their fingertips and use it to make sound, fact-based decisions will organisations be able to use big data to generate operational improvements and profitability that will be hard for competitors to erode. As noted by RP3, RP2, RP5, and RP7, the time it took to implement a big data and analytics initiative resulted in staleness issues and this resulted in the information required for decision making not being as readily at hand as Ross et al. (2013) suggested it should be.

Five participants noted that based on their perspective, leadership buy-in and support for the use of big data and analytics was demonstrated, while the remaining four participants found leadership support and buy-in to be wanting. Vidgen et al. (2017) noted that a point of departure required for a successful big data and analytics journey is a strong, top-down approach steered by leaders who are able to get the commitment of the rest of the firm. Furthermore, Henke et al. (2016) noted that leaders needed to champion big data and analytics in the organisation. Rosseau and McCarthy (2007) supported this and noted that the successful implementation of evidence-based management, in this case led by data, lay in the hands of leadership who embrace the practice to be able to make it work. They noted that to date, managers had been slow in augmenting intuitive decision making with scientific evidence as a routine part of decision making (Rosseau & McCarthy, 2007). The findings of this research indicated that a lack of extensive data use in decision making existed and that consistently strong leadership was not always noted. This was indicative of the slow uptake of data-led decision making in business, congruent with what Rosseau and McCarthy observed in their 2007 study. Research participants, notably RP2 also noted that business areas often had questions that had not historically been successfully addressed by big data and analytics and hence business remained reliant on intuition and gut feel as the key resources in decision making. The observations around scepticism of using big data and analytics, which was raised by three participants (RP3, RP5, RP6), offered insights into the fact that business practitioners tended to be inclined to other methods to support decision making. This finding was disappointing.

A noteworthy finding was that in four of the five organisations where leaders demonstrated strong buy in, big data and analytics was still not used extensively; thus proving that executive support alone is not sufficient to ensure that big data and analytics are adapted. This therefore illustrated support for a comment made by RP3 - that operational level buy was also essential in achieving data-led decision making.

Also noteworthy was that in two instances (RP2 and RP9) strong leadership support was noted at the same time as a lack of leadership understanding, and this indicated that the leaders do not have a clear sense, at a practical level, of what they were supporting in terms of what big data and analytics required of the organisation. Certainly in these cases the researcher noted that an increased level of understanding of the practical implications of big data and analytics should be obtained if any hope of improved use in the organisation is to be achieved. In the other six cases, the barriers around time, financial resources, and scepticism warrant further consideration.

6.2.2 Ask the right questions

The relevant extract from research proposition 1 is reproduced below. It relates to the expectation that the new dynamics of data-led decision making requires organisations to have the ability to ask the right questions.

To create value from big data and analytics:

- Organisations need to know have the ability to **ask the right questions** at the right time;

The findings in relation to this portion of the proposition are contained in Chapter 5.4.1.2. It indicates that although the six of the seven participants who spoke to this topic agreed about the need to follow a different process of asking questions in a big data and analytics world, the discussions were limited in their depth did not focus on the importance of a distinct process of formulating the right questions. One participant (RP1) did not believe a change in approach was required. Two participants (RP9 and RP5) did not voice any views on asking questions at all.

RP2, RP6, RP8, RP3, and RP7 expressed support for the data and analytics team having the latitude to take the lead in asking questions and find novel insights. They all also noted that it was not optimal if analytics teams embarked on the big data and analytics journey in isolation; they needed to work with business. Despite this, they did not emphasise a need to have a process of ensuring that they get to the right questions. This was at odds with the literature reviewed as per Chapter 2. Saran (2015) quoted Debra Logan, the vice president of Gartner, and noted that the required change in leaders' decision making needed to be that they started by asking what business outcomes, for example, new revenue streams, should occur as a result of a project touching on the importance of asking the right question about which problem needed to be solved to get the appropriate answer. Willmott and Dewhurst (2014) supported this role of leaders in asking the right questions at the right time. However, this aspect was missing from the discussions. Although it was touched on by participants, it was not clear that the business leader or the analytics team consistently engaged in a distinct process of asking questions to ensure the right question was identified. The researcher interpreted the lack of detailed responses around making sure the right questions were being asked as indicative of the fact that the participants do not view the ability to ask appropriate questions as a key part of the decision making process in which big data and analytics was used to create value. Instead, they researcher understood that participants interpreted the process of asking questions as part of collaboration.

In terms of where questions should originate, RP6 and RP2 believed that either business or the analytics team can ask a question, while RP1 and RP4 offered a counter view, and believed that it was the role of business to articulate a question. The literature also offered support for originating ideas in business as per RP1 and RP4, as Henke et al. (2016) noted that it was important to have in-house people with the right business skills and corporate knowledge to ask the right questions to identify business problems and translate the results into business solutions. These people operate as partners with the data scientists in creating value. This therefore tied in with the view expressed by RP6 and RP2 that even if an idea is seeded in business, the analytics team should not get stuck in a paradigm whereby it simply takes instructions from business and does not contribute anything to the formation of a business problem through asking questions. This also aligned with the concept of having to co-create as per the preceding paragraph, irrespective of where a question originates. Hence, asking the right question was not distinct from the process of collaboration.

Two participants, both executives, did not comment on the topic of asking questions. Based on the interviews of these individuals, the researcher concluded that although RP5's organisation used data extensively he was not close enough to the operational level of the big data and analytics to offer a view on the process in which questions are defined to be pursued. On the opposite side of the scale, RP9's organisation was in its infancy in using big data and analytics and the processes may not yet have been refined to this level of granularity.

6.2.3 Collaboration

The relevant extract from research proposition 1 is reproduced below. It relates to the expectation that the new dynamics of data-led decision making requires more collaboration.

To create value from big data and analytics:

- Organisations need to have collaboration between their data and analytics functions and business.

According to Daft (2011) leaders need to encourage organisations to collaborate to derive value from big data and analytics. Henke et al. (2016) concurred, and noted that leaders needed to encourage department to work together. As illustrated in Chapter 5.4.1.3, all participants agreed with this, and noted that that creating value from big data and analytics required increased levels of collaboration. This collaboration occurred between big data and analytics teams and the business areas they service; big data and analytics teams and their external clients; and between the various big data and analytics teams within an organisation.

RP6 (39%), RP3 (24%), and RP2 (21%) were the top three contributors on the topic of collaboration. RP6, RP3, and RP2 were also the only three participants who were not executives and, ostensibly, their comments are of a more granular nature as they are closer to the operational level at which big data and analytics initiatives are executed. RP2 noted that collaboration does not result from an instruction that is pushed down from management. Instead it results from the imperative to combine skill sets to deliver value and is therefore the product of ongoing engagement. RP3 also picked up on the idea of partnering with business as a means to help them better understand how their problem had been tackled and resolved, equipping them with "lingo" to be better versed in the big data and analytics conversation. RP6 also noted that participation in business forums brought data and analytics teams closer together. Thus to have greater collaboration between big data and analytics teams and the business areas they service an active drive to create the appropriate levels of engagement needed to be in place. This finding, that collaboration does not simply arise but needs to be nurtured, is aligned with the literature. Henke et al. (2016) noted that leaders needed to overcome resistance to the use of big data and analytics and to encourage departments to work together in leveraging it.

Davenport (2013a) maintained that companies need general managers to partner with data scientists. These general managers will play the critical role of building propositions based on the insights produced by the data scientist and combined with their (the general manager's) knowledge of the business. To be able to partner successfully, general managers need to understand analytics. The general manager and data scientist also need to build trust, so that they are able to ask the data scientist tough questions and also freely swop ideas to solve business, rather than just mathematical problems. RP6 noted that through this close relationship, business started to see the data analytics team and their role differently. He stated that "We go from being analytics to being insights", indicating that the collaboration also creates a greater appreciation for and trust in the big data and analytics teams in an organisation. However, as RP7 stated, big data and analytics teams should prove themselves to business to be able to get this trust and prove that "we can appropriately make sense of the figures that we can add value in your life".

Participants discussed how a single, centralised data and analytics team or multitude of data and analytics teams dispersed throughout the organisation contributed or detracted from collaboration around big data and analytics. The participants' organisational structure represented a fairly even split of centralised compared to decentralised big data analytics teams.

Instead of focusing on which of the two structures is more optimal, RP9 expressed a view that a centralised big data and analytics function should have a neutral reporting line, such as into an executive head of a support area so that it can remain neutral in prioritising and therefore fair in serving the competing needs of various business areas.

A problem that can arise in a decentralised structure is that disparate data sources used internally can produce different results when the same work is performed by different teams – duplication and inefficiencies may arise (RP3, RP5, RP7). Due to these factors, amongst others, RP3 and RP7's organisations had moved to centralised features. This change is congruent with the findings by Ross et al*.* (2013), who noted that the first step was to undergo a disruptive change to concisely apply data-lead, evidence-based decision making. They postulated that an organisation should undergo a disruptive change in organisational structure to accommodate the efficient diffusion of big data and analytics and decision making based on it. However, RP7 and RP5 noted that decentralised functions can also remain highly functional, provided that forums and communities of practice are instituted to deliberately break down barriers so that big data and analytics teams within an organisation could share knowledge and ensure that the insights generated from data are shared. Along the same line of recommendations, RP8, RP4, and RP1 mentioned that for successful big data and analytics project execution, business had to take the lead, based on an understanding of what is possible.

As both a centralised and decentralised approach has its merits and drawbacks, and given that finding a superior structure was not the aim of this research, the researcher was unable to conclude on what organisational structure was best for creating value from big data and analytics. Instead, a key aspect is that business takes the lead in the conversations and that a culture of partnership between business and the analytics teams prevail. The optimal organisational structure to promote the widespread adoption of big data and analytics within an organisation represents a potential area for future study.

Collaboration between big data and analytics teams and the clients of an organisation

RP3, RP9, and RP7 noted that closer collaboration with their corporate clients was also an imperative for the organisation. A reason for these three participants having ventured down this route could be the corporate-to-corporate nature of the relationships in question, with RP9 and RP7 focussing on their agent-principal relationship, and RP3 as a corporate and investment bank (CIB) having greater awareness of its clients. Thus the nature of different clients presents differing opportunities for collaboration. The views expressed by RP9, RP7, and RP3 aligned with the research conducted by Davenport and Kabyla (2016), which indicated that large bodies of data and related analytics can in itself represent digital assets, which can be made available to customers and thus create value. What was surprising is that more leaders did not view data as an asset to be harvested in this manner. The relative lack of maturity in the big data and analytics journey may be accountable for this.

6.2.4 Conclusion

In terms of decision making, it was found that leaders of five of the nine organisations supported the big data and analytics journey, while three leaders did not fully understand what the journey entails. As a result, the researcher questioned the credibility of their buy-in and ability to inspire board-based adoption of big data and analytics. The finding that only one of the nine organisations included in this research used big data and analytics extensively provided support for this scepticism. However, many organisations noted that they were early in the big data and analytics journey, and therefore, on balance it appears that organisations believed in using data in decision making.

Although seven participants were able to discuss the process of asking questions, the researcher was not convinced that they necessarily perceived an increased importance of asking the right questions when exploring business problems to be solved using big data and analytics. In this regard, the researcher noted that the discussion on this topic was more analogous to a discussion on collaboration as the discussion of questioning and refining questions was made in the context of co-creation. Asking the right questions was not a distinct focus area as yet. One potential reason for this lack of focus on finding the right questions could be the fact that big data and analytics were not used extensively. It is possible that only once an organisation used data and analytics consistently and pervasively would the process be refined to the extent that asking the right questions become a distinct exercise.

The extent and depth of engagement of all participants indicated conviction that the big data and analytics journey requires increased collaboration both between analytics teams and business areas they service, but also between analytics teams and the clients of the analytics team, as well as various analytics teams in a decentralised organisation. The researcher also noted that decentralised analytics teams across the organisations appeared to grapple more with collaborating amongst themselves than the analytics teams appeared to grapple with collaborating with the business units they serviced. The participants noted an uptick in collaboration within their organisations, but noted that this did not happen organically and that change management initiatives needed to be instituted to overcome hurdles such as scepticism from business practitioners.

Overall, research proposition 1 was supported as the leaders of organisations demonstrated support for the use of data in decision making. Collaboration was seen as an important part of leveraging big data and analytics to create value, with the process of asking the questions part of collaboration rather than a distinct process.

6.3 Discussion of research proposition 2: Resources

To achieve value out of big data and analytics, organisations need to have the enabling **resources**, which include:

- **technical** skills/capabilities;
- **financial resources** and **infrastructure.**

The two aspects of this proposition are discussed separately in Chapter 6.3.1 and 6.3.2. An overall conclusion for research proposition 2 presented in Chapter 6.3.3.

6.3.1 Skills

This part of research proposition 2 was borne from the literature review which indicated that there were high levels of demand for people who possessed the requisite skill sets to be able find insights through analytics, enabling decision makers to use data on an ongoing, rather than ad-hoc, basis. Henke et al. (2016) noted that the right capabilities were critical to the successful implementation of big data and analytics. In this regard, Henke et al. (2016) referenced data scientists, who had the technical skills to process the big data using meaningful analytics, as being important.

The findings contained in Chapter 5.4.2.1 indicated that leaders recognised that to become data-led, organisations required new skill sets. The majority of participants (six) noted that the critical skill sets were quantitative in nature and used the words data scientist or analyst as collective descriptors of people who embodied these skills. This is consistent with the Davenport and Patil (2012) who referred to such people, who are able to discover new insights while being immersed in data, as data scientists. Furthermore, three participants believed that diversity in skillsets were important, and even though some quantitative skills are required, other backgrounds, such as engineering, could also add value in a big data and analytics world. This is aligned with the view expressed by Daft (2011), who noted that modern leaders valued differences amongst people.

Six participants noted that data scientists needed to possess business acumen or an understanding of the business so that they were able to partner with business to define and solve business problems through big data and analytics. These participants were therefore in agreement with Davenport and Patil (2012), who opined that to have an impact, data scientists should have an understanding of the business and an awareness of customer needs, combined with the ability to communicate proficiently with business around these. One participant noted that the business partner, not the scientist, needed to possess the business acumen. Although this is aligned with the literature noted in Chapter 2 in which Henke et al. (2016) posited that the business acumen would be embedded in a separate individual who would work closely with the data scientist. The researcher observed that this would only work in cases where strong collaboration already existed, rather than where it still needed to be fostered.

Furthermore, as detailed in Chapter 5, three participants noted that there was a shortage of the required data scientist skill sets. This was consistent with the assertion made by Davenport and Patil (2012) who stated that data scientists were rare. This scarcity resulted in the data scientists attracting premium salaries and being difficult to attract and retain.

Three participants noted that the nature of the skill sets required were dynamic and that the staff involved in big data and analytics needed to be willing to update their knowledge to remain relevant. It was also noted that organisations needed to have platforms to facilitate this continuous learning. Combined with the observation made by participants that a diversity in skillsets are valuable, the researcher understood that although it is useful for data scientists to have quantitative ability, their value to an organisation actually came from their ability to apply their business acumen and sense the changing environment and to adapt their skillsets to stay relevant as the world the organisation operated in changed. The researcher did not explore whether an organisation that encouraged and facilitated the evolution of skillsets kept staff more engaged and therefore experienced less attrition.

Overall, the findings outlined above supported the research proposition that the big data and analytics journey requires different skill sets from that which traditionally prevailed within an organisation, however this skillset was not as narrow as simply technical skills as initially proposed, although quantitative skills were important. Business acumen, diversity and an ability to learn continuously are also key skills required to leverage big data and analytics successfully.

6.3.2 Financing and physical infrastructure

Eight participants noted that it was important that an organisation transition to having the appropriate physical infrastructure to support the generation of value through big data and analytics. Three participants opined that the successful implementation of big data and analytics programmes required supporting infrastructure that could consolidate data. Reference was made to non-traditional, cloud-based infrastructure such as data lakes, which are able to accommodate both the volume and velocity of big data to provide real-time data flows. In principle, the findings of the study were aligned with the literature. Seddon and Currie (2017) noted that open-source software, and storage infrastructure became increasingly important in a big data an analytics world. Similarly, Delen and Demirkan (2013) argued that service-oriented, cloud-based resources and infrastructure were required to permit the necessary capability, scalability, yet flexibility to enable the realisation of the value big data and analytics offers. The extent of legacy issues that prevailed in organisations proved to be somewhat surprising as it offered a barrier, compounded by the funding requirements this introduced. In addition, the point was made that implementing the appropriate infrastructure could take a significant amount of time due to the size of an organisation, while outdated organisational policies may not permit the implementation of technology stacks that are required. All of this added complexity to transformation of the IT environment, the pains of which was evident in the discussions.

In terms of financing, four participants noted that the big data and analytics journey required financing – this went beyond simply the infrastructure costs. Three participants noted that traditional return on investment considerations applied when embarking on the big data and analytics journey and that the value potential needed to be demonstrated to secure funding. One participant held a different view and noted that investing in big data and analytics capabilities required a different mind-set in which organisations simply invest in the hope of generating some form of return, which may not be known or quantifiable in practice. This dissenting view was foundationed on that fact that the newness and exploratory nature of the big data and analytics, meant that organisations would not be able to reliably quantify what returns it could generate from a big data initiative. Hence any traditional means applied to measure a return on investment may be unsupportable by experience and therefore flawed. Barton and Court (2012) posited that leaders were weary of making significant investments to position their organisations to take advantage of big data and analytics as they were uncertain of how to proceed to exploit big data and analytics. The research finding that stakeholders required a return of investment to be demonstrated upfront is indicative that organisations are weary of making investments that may not generate returns and used traditional investment deaccessioning methods as a way of managing some of that uncertainty. However, a case is made that stakeholders should move away from this mentality and view big data and analytics as a prospecting opportunity. The view is also supported by the findings on value, as discussed in Chapter 6.4, which indicated that participants did not identify metrics which could be used to demonstrate the value that was generated through big data and analytics. As a result, a traditional financial investment model would have the same shortcoming.

Based on the above, the findings support the research proposition that IT infrastructure and financing resources are important additional requirements for successful creation of value from big data and analytics.

6.3.3 Conclusion

Based on the research findings as discussed above, research proposition 2 is generally supported.

6.4 Discussion of research proposition 3: Value

To create value with big data and analytics, organisations need to be able to **articulate** what **value** it is looking for, as this enables articulation and tracking of success **metrics**.

If a company wishes to survive, it is critical that it creates value when there are fundamental changes in what companies analyse, create, and deliver through big data and analytics (Lindgreen et al., 2011). To be able to create value, rather than just undertaking endless experiments of questionable value, organisations needed to be able to clearly specify how it will be used to generate value and how the results will be measured (Henke et al., 2016).

The findings contained in Chapter 5.4.3 indicated that leaders conceptualised the value to be derived from big data and analytics in four broad categories. The research findings of what constitutes value were aligned with the literature in the following ways:

- McAfee and Brynjolfsson (2012) suggested that factual financial and operational metrics could be used to define value. They found that organisations that had taken the lead in their industry in using data-driven decision making achieved 5% higher productivity and 6% higher profitability outcomes as compared to their competitors, after controlling for other potential contributors to the results, such as labour. The participants in the research study believed that financial (67%) and operational (11%) results can be achieved as a result of leveraging big data and analytics in their organisation. However, they were unable to quantify the extent of the value contributed to date or to articulate an expected quantum of value.
- Kiron and Shockley (2011) explored the generation of competitive advantage as value and found that 57% of participants to their survey were gaining competitive value from big data and analytics. Although no participant claimed to have necessarily realised competitive value from big data and analytics as yet, 33% of participants in this research believed that they could achieve a competitive advantage from big data and analytics. Once again, none of the participants articulated precisely how this was being tracked and measured.
- Davenport and Kabyla (2016) posited that large bodies of data and related analytics, can in itself represent digital assets, which could be made available to customers and thus create value directly and as a revenue source. Two participants (RP1, RP2) supported the view that their organisation’s big data and/or analytics are in and of itself assets which were, or could be, sold to generate value for the organisation. One participant (RP9) contended that their organisation did not believe that they were able to monetise their big data and analytics as a good or service due to the adverse impression this would create with their clients. This presented an interesting data point for organisations who have other business organisations as customers,

as it suggested that organisations should consider the potential trade-offs between aspects of customer relationship management and experience compared to a direct monetary realisation of value from big data and analytics.

- Wamba et al. (2016) referred to value as improved efficiency. This is consistent with what RP9 defined as their organisation's conceptualisation of value.

However, none of the participants demonstrated the ability to clearly articulate what the quantum of that value would be. RP2 offered an explanation for this and noted that an inability to isolate the results of actual big data and analytics projects meant that the organisation was unable to demonstrate if it had truly achieved success. The researcher therefore noted that metrics needed to be agreed upfront to enable tracking of value, and this did not appear to have been done in any of the organisations represented in this study. RP7 noted that a lack of actual delivery of big data and analytics projects also resulted in an inability to demonstrate quantifiable value.

6.4.1 Conclusion: Value

The participants did not have consensus around what value would be, with different participants conceptualising it as financial impacts (67%), improved competitive advantage (33%), data and analytics as a good or service (22%) and operational efficiencies (11%). Moreover, although the conceptualisation is in place, none of the participants clearly articulated the quantum of value they were pursuing. Thus, they were unable to demonstrate whether data and analytics had been successful in delivering value. The contention made in research proposition 3 that organisations had value measures in place was therefore not supported by the findings – a conceptualisation of value appeared to be sufficient in the early stages of a big data and analytics journey.

6.5 Discussion of other findings: Maturity in big data and analytics

Getting the basics right: data

Delen and Demirkan (2013) noted that the effectiveness of analytics depends on the quality and quantity of data, the integrity of its management, and the sophistication of analytical tools used to process the data. They noted that as organisations became more sophisticated, they moved beyond looking at understanding the past and current occurrences through descriptive analytics, and moved toward exploring what will happen in the future, using predictive analytics. At the most advanced level, organisations would use prescriptive analytics, and this was where analytics will determine what actions organisations should undertake. Based on the description of how analytics are used in organisations included Chapter 5.5, the application was mostly descriptive, with some organisations tending toward predictive use of analytics.

This surprised the researcher, whose expectation was that the financial services sector would be more advanced in the use of big data and analytics.

Boyd and Crawford (2011) also noted that large data sets came with inherent issues which are often not acknowledged. Data quality became an issue as large data sets often contained weaknesses, such as errors, gaps, and sampling biases which were exacerbated by combining various data sets and hence, it was important to understand the providence of data. Based on the views expressed by participants, it was these issues that are delaying their evolution to using more predictive and prescriptive analytics, as organisations first needed to sort out their data issues. For example, RP1 noted that a reason organisations are focussed on MI and BI is that they are laying the foundations for the next steps.

> *"So I don't think you have to be perfect at your BI and MI to do any form of analytics. I'm just saying* ***without the basics in place, it's hard to make that jump holistically****."*

Willmott and Dewhurst (2014) noted that big data differed from traditional data in how it was aggregated and managed. The efforts of participants in the study to land their data are testimony to that. Participants noted particular focus on consolidating and centralisation of data. Saran (2015) called for a change in the approach of data governance through the democratisation of data within organisations, centralising that which is required to be kept private due, while decentralising the storage of that which can be shared freely– encouraging the wide spread use of big data. Based on the view expressed by participants, their priority was not on empowering business with data and democratising it, but rather to gain control of and reconcile the data. Data quality was their main focus.

Willmott and Dewhurst (2014) made the case for not only internal, but also external data sources to be combined to create valuable big data sources. The findings around the use of unstructured and external data being in its nascent stages dovetails with the research findings by DalleMule and Davenport (2017), who found that less than 1% of unstructured data are used by organisations to inform business decisions. The focus by organisations on "getting the basics right first" are also aligned with Ross et al*.* (2013), who posited that most organisations were not yet able to use traditional data effectively and hence, despite huge investments in big data and analytic capabilities, they were not able to create value from big data. The participants of the research provided evidence to this.

Based on the focus time and effort exerted by participants on getting their data sorted and first exploiting structured data fully before embarking an extensive big data and analytics journey, indicates that having data in order should have been included in the conceptual model presented in Chapter 3 as a precursor to being able to create value from big data and analytics.

Qualitative methodologies

Boyd and Crawford (2011) and Crawford (2013) espoused that big data should be combined with qualitative techniques to add greater depth and more meaning to big data, as qualitative information provides context to big data. Based on the literature presented by Wang (2016), it appeared that the use of qualitative methodologies would be a logical extension of a big data and analytics program in which the consumer is at the centre. Many financial institutions included in this study professed customer experience to be a key concept of value and hence, it was considered that qualitative analytics could we applied in decision making. The use of qualitative methods was therefore not explicitly included in the model, in a similar way that analytics itself was not explicitly included in the model, but a foregone conclusion.

The findings of this research indicated that only RP's organisation made use of qualitative methodologies. The remaining participants used only quantitative analytics, and this called into question the proclamations made by participants that they pursued improved consumer experiences as part of the value to be derived from big data and analytics. This is because thick, qualitative data offered an opportunity to gain unique insights into consumers that would not be revealed by big data an isolation, but this opportunity was not ceased. The lack of maturity in using predictive analytics and the use of mainly structured data could be used as an explanation of this, as qualitative techniques could be seen as the next step in the journey. However, given that qualitative methods are stand-alone, and can be used irrespective of the quantitative journey – as RP6's organisation has gone some way toward proving – the researcher does not believe speculation is appropriate and would suggest that this is an interesting area for further research.

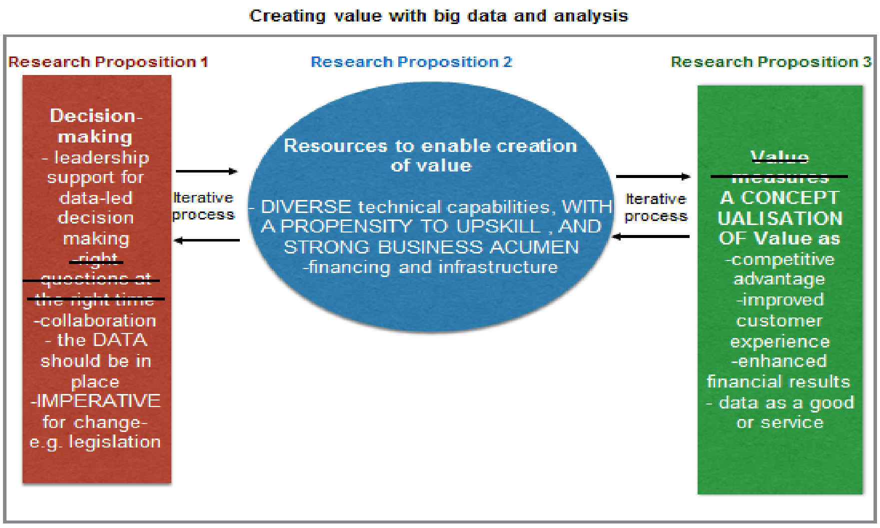
Adoption challenge

Each organisation seems to have realised the need to actively drive adoption of big data and analytics and that it does not necessarily sell itself to an organisation. The conceptual model proposed in Chapter 3 did not include the need for an imperative. However, based on the conviction case made by participants, demonstrating that an organisation needs to have an imperative for change, the researcher acknowledges this as a shortfall of the initially proposed model, and that it should be updated accordingly. This also presents an area of further study.

6.6 Implications for the conceptual model

Based on the research findings noted in this chapter, the conceptual model proposed by the researcher in Chapter 2 is only partially valid. The research findings presented a compelling case for the following revisions to be made to the conceptual model, as illustrated in CAPITAL print and stricken-through text in Figure 6.6.1.

Figure 6.6.1: Implications for the conceptual model



- Research proposition 1 should be updated to remove the process of the right asking questions as being a distinct aspect of the process of data-led decision making. The proposition should also be updated to incorporate a "get the basics right first" element. This is formulated as having data in place as this is the element most discussed by users. A further addition to the model is the requirement of having an imperative in place to create momentum in the journey to becoming data driven.
- Research proposition 2, related to the resources required to create value, should be updated to include a requirement to have not only technical skills, but a diversity of skills as embodied in individuals who are driven to continuously educate themselves to stay relevant. Research proposition 2 should also be updated to emphasise the importance of business acumen.
- Research proposition 3 should be updated to include that a conceptualisation of value is required for organisations to embark on the big data and analytics journey, but that clear metrics of success do not need to be predefined upfront, although the researcher acknowledges that this may pose challenges with demonstration of success later on in the journey.

Chapter 7: Conclusion and recommendations

7.1 Introduction

In this chapter, the findings of Chapter 5 and the discussions in relation to the theory in Chapter 6 are consolidated into a framework that intends to summarise the outcomes of this research. A discussion of the implications of the framework for management, limitations of this research and suggestions for possible avenues of future research follow thereafter.

7.2 Key findings

7.2.1 Summary of research findings

The main objective of the research was to discover the changes an organisation undertook in its quest to create value from big data and analytics. The study also identified challenges that arose in this process.

Financial service institutions have a wealth of data and are ideally placed to take advantage of this big data through analytics and thus generate increased value for their stakeholders. Financial service organisations that have adopted big data and analytics have shown that even though it was not pervasively used throughout their organisation to inform decision making, certain changes needed to have been undertaken to accommodate the big data and analytics agenda. There needed to be visible support by leaders who understood the power and potential of big data and analytics. This was not necessarily sufficient to result in wide-spread adoption of data-led decision making. Organisations needed to have an imperative to drive the adoption of big data and analytics. Collaboration between analytics teams, analytics teams and business area, and analytics teams and the client of the organisation also increased and change management initiatives were required to support this change. To be able to successfully exploit the data, organisations need to ensure that their foundations are set – that it has gotten the basics right first in terms of data quality, lineage, and storage. This implied that resources were required – notably financial and infrastructure resources, but legacy infrastructure, outdated policies and traditional means of approaching the investment decisions offered challenges in making a transition. These were compounded by the size of the organisation, which hampered its agility. Moreover, a big data and analytics journey required a different set of skills, with the data scientist playing a key role. Data scientists must not only possess quantitative capabilities, but must also have strong business acumen to allow them to engage with business and clients. Given the rapidly evolving landscape, they also needed to be driven to continuously update their skill set. Finally, organisations needed to know what they were aiming for in using big data and analytics and therefore need to have some conceptualisation of value. Even though the precise quantum of this value of may not be known, organisations could, and did still, pursue big data and analytics if it believed that it could deliver value.

7.2.2 Revised conceptual model

Financial services organisations are data rich and increasingly realising the need to harness that data to improve the goods and services offered to their customers, and develop new products and services based on the unique needs of their customer. The question these organisations are faced with is: how would they go about being able to exploit big data through analytics to be able to create value?

Figure 7.2.2.1: Revised conceptual model

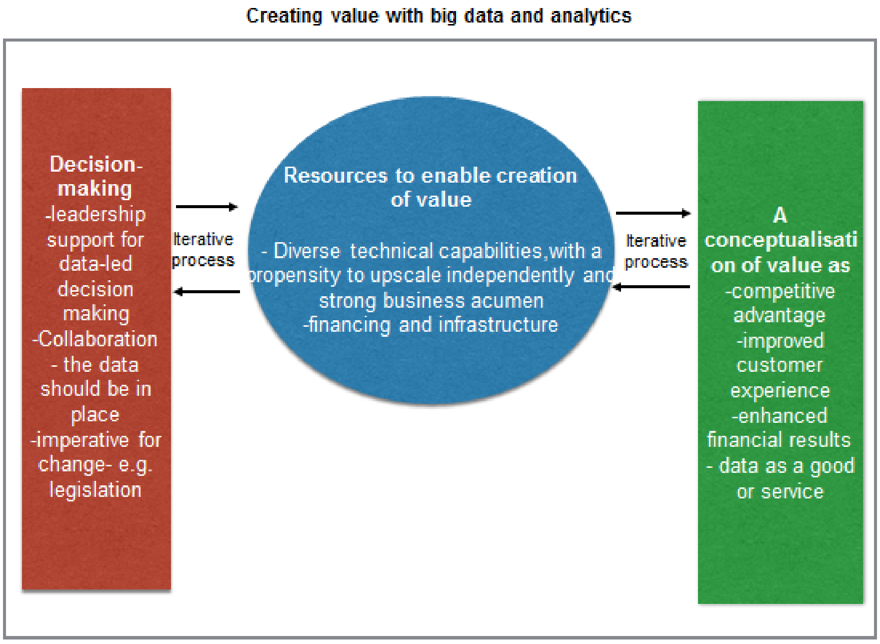


Figure 7.2.2.1 presents a framework that may assist in addressing this question. This figure is based on the conceptual model included in Figure 6.6.1, which was updated to incorporate and illustrate the changes required to the original conceptual framework presented in Chapter 3. Financial services organisations that are about to commence or are in the early phases of adopting big data may use this framework to guide their thinking regarding the changes in approaches to decision making, resources and conceptualisation of value that will be required.

7.3 Limitations of this research

In this section, the limitations of the research based on its execution, are discussed.

Researcher bias

Exploratory research is subjective and is influenced by the researcher's own perspectives. As a result, it is imperative that a researcher recognises and documents those potential biases, as it would have influenced on how the researcher interprets the findings of the research (Creswell, 2014). During the currency of this study, the researcher was employed by a bank which forms part of larger financial services group. This may have had an adverse impact on how the participants responded, potentially biasing them to being less forthcoming with information that they may otherwise have shared, due to concerns that proprietary information may be revealed to an employee of a competitor organisation.

Sampling bias

Four of the participants in this research were employed by the group of financial services organisations in which the researcher is employed. The various organisations and segments within the group operate as separate brands with discrete strategies, boards, and executives directors. A strong owner-managed culture exists and as a result each segment is effectively operating as a separate business. Based on this operating model, the organisations were included within the sample as they offered different perspectives. However, the researcher does acknowledge that this may introduce an element of sampling bias toward the one organisational group.

Participants' influence on the research

Six of the nine participants had a strong quantitative or statistical background and this may tilt the findings toward a more analytical bias than what a more diverse group of organisational leaders with other background may have offered. Six of the nine participants were employed at an executive management level within their organisation and three were in middle management. This resulted in the study being biased toward leaders with a strategic orientation, rather than an operational orientation.

Sampling method's influence on the research

Due to the use of convenience sampling, the majority of the interviews were conducted with participants from the banking sector. As a result, the sample exhibits a bias toward the banking sector as a subset of the financial service industry. The prominence of the banking sector means that findings may not be relevant to other areas and should not be extended beyond this sector (Seddon & Currie, 2007). The convenience sampling also meant that the relationships (such as friendships, or business relationships) which lead to the identification of the participants could have resulted in them being similar in certain personality traits, beliefs and experiences, introducing a further bias into the research.

7.4 Implications for management

This research has identified a number of implications for the management of financial institutions. Notably, big data and analytics adoption is not a “plug-in” solution and has wider implications that require consideration by management. Managers need to consider the following factors:

- The level of understanding the organisation’s leaders have around big data and analytics and its implications.
- The level of buy-in to evidence-based decision making within the organisation.
- The extent to which it is using its current structured and unstructured data.
- The level of data readiness – the sources, of data, the completeness of data, the extent of consolidation of data and its quality are amongst the things to be considered to determine if the organisation has got the basics in place to be able to advance to the big data and analytics stage.
- The extent to which the organisation operates in silos, or is able work together.
- The existence of an imperative for change – be it legislative or other – that can be leveraged as a means to springboard the data and analytics journey from.
- The diversity of the technical skill set and extent of business acumen and drive for continuous learning embodied in the data scientists, as well as the strategies for retention and motivation of these employees.
- The required infrastructure to be able to support the data and analytics journey and the funding required for this as well as other resources required to create value from big data and analytics.
- The existence of a conceptualisation of the value the organisation wishes to realise through the big data and analytics journey.

If the organisation has clarity on where it stands in relation to the above considerations, it will be better placed to gauge the level of success it is likely to achieve in creating value from big data and analytics.

7.5 Recommendations for future research

This research established a framework for successful creation of value through big data and analytics in financial services organisations. It is recommended that as a next step, research is undertaken to test that framework with a larger population of organisations.

As noted in Chapter 6, organisational structure was a topic that was top-of-mind for many participants. As a result, it presented a case for further research into the optimal organisation of analytics teams as centralised or decentralised within an organisation to best enable it to exploit big data and analytics and to embed a big data culture.

A finding of this research was that organisations needed to have an imperative for using big data and analytics to ensure it was adopted throughout the organisation and seamlessly integrated into decision making processes. However, what these imperatives may be was not fully explored within this research, and therefore this also presents an area for further study.

Organisations also did not make use of qualitative methodologies. Exploring why qualitative methodologies were not used, despite its relevance in providing a deeper understanding of context, presents an area for further research.

7.6 Conclusion

This research explored three research propositions in relation to the process, as perceived by leaders of financial service organisations, through which big data and analytics was used to create value. The first proposition explored the extent to which leaders of organisations demonstrated support for data-led decision making, and how the decision making process had changed to include a robust process of asking the right questions and greater collaboration in a data-led world. Secondly, it explored the extent to which different resources were used in the decision making process, focussing on the skills as well as the physical and financing demands created by a big data and analytics journey. Lastly, the research explored the extent to which organisations that made use of big data and analytics were able to conceptualise and clearly demonstrate the value they were creating through their processes.

The findings of this research have demonstrated that the leaders of organisations demonstrate support for the big data and analytics journey. However, evidence-based decision making was not used extensively within organisations, and potentially as a result of this, the process for asking and refining the right questions has not an area of distinct focus. Organisations needed an imperative to drive the adoption of big data and analytics- this change did not happen automatically. Despite this, collaboration between analytics teams, analytics teams and the business units they service, as well as between analytics teams and the clients of the organisation had improved. The skill sets required were technical in nature, as embodied in data scientists, but these skills needed to be diverse, and embodied in individuals who were driven to learn on a continuous basis. It was also important that data scientists had strong business acumen, which not only enabled them to be able to remain relevant, but also enabled good collaboration. The big data and analytics journey also required additional financial and infrastructure resources, although outdated corporate policies and the size of the organisation posed barriers to getting the right resources. Finally, it was found that although an organisation could conceptualise what value it was pursuing through big data, it was unable to articulate and demonstrate the quantum of value clearly, for example as a 10% increase in performance. This did not hinder an organisation from pursuing big data, but it did introduce complexity in gauging whether the objectives of a big data and analytics world was achieved.

Appendices

Appendix 1: Example of consent form

Informed Consent

I am conducting research on leadership challenges faced in the use of big data and analytics to create value. Our interview is expected to last about an hour, and will help us understand the challenges organisations face in creating value from big data and analytics and potential methods of overcoming these challenges. **Your participation in this research is voluntary and you can withdraw at any time without penalty**. Of course, all data will be reported anonymously in the research i.e. without identifying you. This interview will be audio recorded. If you have any concerns, please contact my supervisor or me. Our details are provided below.

Researcher name: Melissa Anine Stevens
Email: 16391994@mygibs.co.za
Phone: [redacted]

Supervisor name: Robert Beney
Email: Robert@ironsky.co.za
Phone: [redacted]

Signature of participant: ___________________________________
Date: _________________

Signature of researcher: ___________________________________
Date: _________________

Appendix 2: Discussion guide

Initial discussion guide

Grand tour question

- In your view, how extensively do you use data and analytics?

Further prompting questions

- What processes do you use to ensure you ask the right questions?
- What process do you use to ensure you produce credible analytics?
- Tell me how about the role leadership has played in your big data and analytics journey?
- Tell me about the decision making culture of the organisation?
- Lets talk about collaboration and how that may have changed, if at all?

Closing

- What other challenges that we have not spoken about do you face in leveraging big data and analytics?

Final discussion guide

Grand tour question

- In your view, how extensively do you use data and analytics?

Decision making

- How do you ask the right questions?
- What types and sources of data do you use?
- What challenges do you encounter in implementing analytics?
- What is the stance of leadership around the use of big data and analytics?
- Tell me about the level of collaboration?

Value

- How would you define the value generated from big data and analytics?
- Tell me about your ability to generate value from big data and analytics

Resources

- What are your key resources that enable the use of big data and analytics?

Closing

- What other challenges that we have not spoken about do you face in leveraging big data and analytics?

Appendix 3: Themes and phrases that guided the coding process

Tie back to research proposition (RP)	Themes	Codes or phrases of text
Decision making	Leadership	buy-in, vision, leadership support, forward thinking, guidance, taking the lead, top-down, from the top, senior management, top management, ExCo, direction, steer, control, tell, command, governance, supervision
Decision making	Asking the right questions	problem statements, use case, problem identification, questions, queries, enquiry, challenge, robust discussion
Decision making	Collaboration	team work, working together, cross-function, cooperation, coordination, knowledge sharing
Resources	Data scientist workforce	talent, skills, knowledge, qualification, aptitude, ability, expertise, capability, expertness, competence, aptitude, adeptness, prowess, strength
Resources	Infrastructure	data warehouse, storage, processors, servers, clusters, mainframe
Resourcing	Financing	Funding, money, investment
Resources	Value	Return, benefit, output, metrics, deliverable, measures, KPIs, accountability, returns,

Appendix 4: Ethics clearance

A copy of the ethics clearance letter is included below.

Gordon Institute of Business Science
University of Pretoria

22 June 2017

Melissa Stevens

Dear Melissa Stevens,

Please be advised that your application for Ethical Clearance has been approved.

You are therefore allowed to continue collecting your data.

We wish you everything of the best for the rest of the project.

Kind Regards

GIBS MBA Research Ethical Clearance Committee

Gordon Institute of Business Science
Reg. No. 99/19816/08

26 Melville Road, Illovo, Johannesburg
PO Box 787602, Sandton, 2146, South Africa

telephone **(+27) 11 771 4000**
fax **(+27) 11 771 4177**

website **gibs.co.za**
University of Pretoria

Reference list

Baller, S., Dutta, S. & Lanvin, B. (2016). *The global information technology report 2016*. Geneva World Economic Forum and Insead. Retrieved from: https://www.weforum.org/reports/the-global-information-technology-report-2016.

Barton, D. & Court, D. (2012). Making advanced analytics work for you. *Harvard Business Review, 90* (10), 78-83.

Bell, P. C. (2015). Sustaining an analytics advantage. *MIT Sloan Management Review, 2017.*

Biehn, B. (2013). The missing vs in big data: viability and value. *Wired Magazine*. Retrieved from: https://www.wired.com/insights/2013/05/the-missing-vs-in-big-data-viability-and-value/.

Boyd, D. & Crawford, K. (2011). Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Information, communication and Society, 15*(5), 662-679.

Briner, R. B., Denyer, D. & Rousseau, D. M. (2009). Evidence-based management: Concept clean-up time? *Academy of Management Perspectives*, *23*(4), 19–33. http://doi.org/10.5465/AMP.2009.45590138.

Brown, B., Court, D. & Willmott, P. (2013, November). Mobilising your c-suite for big-data analytics. Retrieved from: https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights.

Crawford, K. (2013). The hidden biases in big data. *Harvard Business Review*. Retrieved from https://hbr.org/2013/04/the-hidden-biases-in-big-data.

Chen, H., Chiang, R. H. L. & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4), 1165–1188.

Citi Business Advisory Services. (2015). *Big data and investment management: the potential to quantify traditionally qualitative factors*. Citigroup. Retrieved from: https://www.cmegroup.com/education/files/big-data-investment-management-the-potential-to-quantify-traditionally-qualitative-factors.pdf

Corte-Real, N., Oliveira, T. & Ruivo, P. (2016). Assessing business value of big data analytics in European firms. *Journal of Business Research, 70,* 379-390. https://doi.org/10.1016/j.jbusres.2016.08.01.1

Creswell, J.W. (2012). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (4th ed.). University of Nebraska Lincoln: Pearson.

Creswell, J. W. & Miller, D.L. (2000). Determining validity in qualitative inquiry. *Theory into Practice, 39*(3), 124-130.

Creswell, J.W. & Miller, G. A. (1997). Research methodologies and the doctoral process. *New Direction for Higher Education, 99,* 33-46.

Daft, R.L. (2011). Chapter 1: What does it mean to be a leader? *The leadership experience.* Stamford, CT: South-Western Cengage Learning. Retrieved from: GIBS Blackboard: http://gibs.blackboard.com.

DalleMule, L. & Davenport, T. H. (2017). What's your data strategy? *Harvard Business Review. 95*(3), 112-*121.*

Davenport, T. H. (2013a). Analytics 3.0. *Harvard Business Review*, *91*(12), 64-72.

Davenport, T. H. (2013b). Keep up with your quants. *Harvard Business Review*, *91*(7/8),120-123.

Davenport, T.H (2014). *How strategists use big data to support internal business decisions, discovery and production. 42,* 45-50. DOI 10.1108/SL-05-2014-0034.

Davenport, T. H. & Kudyba, S. (2016). Designing and developing analytics-based data products. *MIT Sloan Management Review*, *58*(1), 82-89.

Davenport, T. H. & Patil, D. J. (2012). Data scientist: the sexiest job of the 21st century. *Harvard Business Review, 90*(10) 70-76.

Delen, D. & Demirkan H. (2013). Data, information and analytics as services. *Decision Support Systems. 55,* 359-363. doi:10.1016/j.dss.2012.05.044.

George, G., Haas, M. R. & Pentland, A. (2014). *Academy of Management Journal, 57*(2), 321-326. DOI: 10.5465/amj.2014.4002.

George, G., Osinga, E. O., Lavie, D., & Scott, B.A. (2016). Big data and data science methods for management research. *Academy of Management Journal, 59*(5), 1493-1507. http://dx.doi.org/10.5465/amj.2016.4005.

Gil, D. & Song, I. (2016). Modelling and Management of big data: challenges and opportunities. *Future Generation Computer Systems, 63,* 96-66. http://dx.doi.org/10.1016/future.2015.07.019.

Henke, N., Bughin J. & Chui, M. (2016). Most industries are nowhere close to realising the potential of analytics. *Harvard Business Review digital articles, Dec 2016 2-4. Harvard Business Review.*

Kiron, D. & Shockley, R. (2011). Creating business value with analytics. *MIT Sloan Management Review, 53*(1), 57-63.

Lavalle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Special report: analytics and the new path to value. *MIT Sloan Management Review, 52*(2), 22-32.

Learmonth, M. (2006). Is there such a thing as evidence based management? A commentary on Rousseau's 2005 presidential address. *Academy of Management Review, 31*(4), 1089-1091. http://doi.org/10.5465/AMR.2006.22528177.

Lindgreen, A., Hingley, M. K., Grant, D. B. & Morgan, R. E. (2012). Value in business and industrial marketing: past, present and future. *Industrial Marketing Management, 41*(1), 207-214.

Lugmayr, A., Stockleben, B., Scheib, C. & Mailaparampil, M. A. (2017). Cognitive big data: survey and review of big data research and its implications. What is really "new" in big data? *Journal of Knowledge Management, 21*(1), 197-212. http://dx.doi.org/10.1108/JKM-07-2016-0307.

Mason, M. (2010). Sample size and saturation in PHD studies using qualitative interviews. *Forum: Qualitative Social Research, 11*(3). Retrieved from: http://www.qualitative-research.net/index.php/fqs/article/view/1428/3028.

McAfee, A. & Brynjolfsson, E. (2012). Big data: the management revolution. *Harvard Business Review, 90*(10), 60-68.

McCracken, G. (1988). *The long interview*. Newbury Park, California: Sage Publications.

Mortensen, M. J., Doherty, N. F. & Robinson, S. (2014). Operational research from taylorism to terabytes: a research agenda for the analytics age. *European Journal of Operational Research*, *241*, 583-595 http://dx.doi.org/10.1016/j.ejor.2014.08 .029

Ndlela, L. T., & du Toit, A. S. A. (2001). Establishing a knowledge management programme for competitive advantage in an enterprise. *International Journal of Information Management*, *21*(2), 151–165. http://doi.org/10.1016/S0268-4012(01)00007-X.

Oliver, D. G., Serovich, J. M., & Mason, T. L. (2005). Constraints and opportunities with interview transcription: towards reflection in qualitative research. *Social Forces. 84*(2), 1273-1289.

Plessis, M. D., & Boon, J. A. (2004). Knowledge management in eBusiness and customer relationship management: South African case study findings. *International Journal of Information Management*, *24*(1), 73–86. http://doi.org/10.1016/j.ijinfomgt.2003.10.002.

Raelin, J. (2015). Rethinking leadership. *MIT Sloan Management Review, 56*(4), 96-97.

Ranyard, J. C., Fildes, R. & Hu, T. (2015). Reassessing the scope of OR practice: the influences of problem structuring methods and the analytics movement. *European Journal of Operational Research*, *245*, 1-13. DOI: 10.1016/j.ejor.2015.01.058.

Ross, J. W., Beath, C. M. & Quaadgras, A. (2013). You may not need big data after all. *Harvard Business Review, 91*(12)*, 90-98.*

Rosseau, D. M. & McCarthy, S. (2007). Educating managers from an evidence-based perspective. *Academy of management learning and education. 6*(1), 84-101.

Saran, C. (2015). *CIOs are ill prepared for data-driven business and must change approach.* computerweekly.com. 24-30 March 2015, 8-9.

Saunders, M., & Lewis, P. (2012). *Doing research in business and management*. Edinburgh Gate: Pearson.

Seddon, J.J.J. M., Currie, W. L. (2017). A model for unpacking big data analytics in high-frequency trading. *Journal of Business Research, 70,* 300-307. http://dx.doi.org/10.1016/j.jbusres.2016.08.003

Shenton, A. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for information,* 22, 63-75.

Sinkovics, R., Penz, E., & Ghauri, P.N. (2008). Enhancing the Trustworthiness of Qualitative Research in International Business. *Management International Review, 48*(6), 689-713.

Spiggle, S. (1994). Analysis and interpretation of qualitative data in consumer research. *Journal of Consumer Research, 21*(3), 491-503

Thompson, C. (2013). Relying on algorithms and bots can be really, really dangerous. *Wired Magazine*. Retrieved from: https://www.wired.com/2013/03/clive-thompson-2104/. Accessed on 12 July 2017.

Vidgen, R., Shaw, S. & Grant, B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research, 261,* 626-639. http://dx.doi.org/10.1016/j.ejor.2017.02.023.

Wamba, S.F., Gunesekaran, A., Akter, S., Ren, S. J., Dubey, R. & Childe, S. J. (2016). Big data analytics and firm performance: effects of dynamic capabilities. *Journal of Business Research*, 70, 356-365. https://doi.org/10.1016/j.jbusres.2016.08.-009.

Wang, T. (2016). Why big data needs thick data. Retrieved from: https-://medium.com/ethnography-matters/why-big-data-needs-thick-data-b4b3e75e3d7

Willmott, P., & Dewhurst, M. (2014). Manager and machine: the new leadership equation. McKinsey Quarterly. 3(3), 76-83.