Value Oriented Big Data Strategy: Analysis & Case Study

Khaled Himmi*, Jonathan Arcondara†, Peiqing Guan‡, Wei Zhou‡

*Department of Information and Operations Management, ESCP Europe, Paris, France
†Department of Accounting, University of Florida, Gainesville, FL, U.S.
‡Corresponding Author: wzhou@escpeurope.eu

Abstract

Big data is emerging in recent years as an evolutionary phenomenon. Many new concepts and business models driven by data are introduced as a result. We in this research are motivated to investigate the value side of big data. We examine the financial statements in CAC40 companies in order to discover the relationship between stock performance and depth of corporate data involvement. Our results are surprisingly two-folded. There are companies with strong data capability that succeed in the stock market. There are also companies without much data depth that also perform well. Moreover, the result doesn’t show link between the not-well performed companies and the lack of data capability. To decode this surprising result, we reexamine the existing strategic big data literature and discover the missing puzzle pieces. We thus propose a new strategic model that considers both supply chain decision dynamics and data capability. We explain this model based on an airline industry’s case study. We draw managerial implications to conclude this paper.

1. Introduction

The term of Big Data started in the late 20th century and today it represents a wide spectrum of applications and huge potential to improve and create new business opportunities (Chen et al. 2012). It is often linked with the increase in the number of real-time data collected mainly thanks to the boom of social media (Facebook, Twitter, Instagram, etc) and IoT applications (RFID, NFC, smart payment, etc). These new sources of data represent a huge potential for tomorrow’s companies. However, when we look on the number of companies that actually use big data, it is relatively limited for the moment.

Up to now there is no commonly accepted definition of big data. Based on a study led by IBM, several commonly used terms have been used to contribute to the family of big data: a greater scope of information, new kinds of data and analysis, real-time information, data influx from new technologies, non-traditional forms of media, large volumes of data, the latest buzzword, and social media data, among many others. Today, we creates huge amount of information from electronic everywhere: clickstream data, online shopping paths, social network data (Facebook, Instagram, and twitter), and sensor network data (the Internet of Things for instance), etc. Big Data is very often characterized by “the three Vs”: Volume, Variety, and Velocity. Volume is an obvious reference of the mass quantities of data that a company has to analyses to improve its business. Variety is about the complexity of multiple data types, which can be structured, semi structured or unstructured. Velocity is important if managers need to analyze the data streams and make decisions quickly. Traditional data analytics are able to process large volume of data with quite complexed structure. Velocity makes one of the key advantages of today’s big data system that a traditional system is not able to compete. Nowadays we often add a fourth “V” as Veracity, which stands for uncertainty that may come from both internal and external environments, such as economic factors, weather conditions, etc.

Recently many companies set big data as a business priority (Sanders 2014). Driven by data, some common goals include targeting customer choices, forecasting new tendencies, optimizing operations, introducing fancier new products, etc (Brown et al. 2011). The cost is high. Every day, people create 2.5 quintillion bytes of data, enough to fill 10 million Blu-ray discs1. In every 60 seconds, 72 hours of footage are uploaded to YouTube, 216 000 Instagram are posted, and 204 000 000 emails are sent. 90% of today’s data has been created in the last 2 years. Among those data, 90% are unstructured. Those numbers give us an insight of the

1 « Bring Big Data to The Enterprises », IBM Softwares, weblink: https://www-01.ibm.com/software/data/bigdata/what-is-big-data.html
potential value of the industry if big data is correctly used. Can we use a basic economic equation to evaluate the benefit and its associated cost?

We want to know whether big data already has a real impact on today’s companies and whether this impact can be justified by existing models. As shown in Figure 1, we add the fifth “V” of value in the existing descriptive model (LaValle et al. 2011). Value has meanings in two folds. Firstly, data creates value for the firms. Secondly, a firm’s value can be reflected by depth of data usage. Most existing works have attempted to study the prior; we are motivated to investigate the second meaning of value brought up by data in this research.

![Figure 1 The 5V of Big Data](image1.png)

We analyze the financial statements in CAC 40 companies in order to discover the relationship between stock performance and depth of corporate data involvement. Our results are surprisingly two-folded. There are companies with strong data capability that succeed in the stock market. There are also companies without much data depth that also perform well. Moreover, the result doesn’t show link between the not-well performed companies and the lack of data capability. To decode this surprising result, we reexamine the existing strategic big data literature and discover the missing puzzle pieces. We thus propose a new strategic model that considers both supply chain decision dynamics and data capability. We explain this model based on an airline industry’s case study.

The remainder of this paper is organized as follows. In section II, We investigate the integration of data of the CAC40 companies and their stock performance. We discuss strategic implication in Section III. We present a case study to illustrate the strategic models in section IV. We the make concluding remarks and give guidance to future research in Section V.

2. Evidence from Stock Performance

Traditionally most managerial decisions are made based on well-organized and structured data, and analyses are always run on data or on reporting data that are already aggregated. On the opposite, the big data approach is often based on the low level data with all the data being collected and analyzed “as is” because of the difficulty and associated cost to manipulate data in such a large volume and variety.

Another shift data analysis in many companies is from search for causality to pattern discovery. Unlike traditional data analysis that often involves hypothesis, which has to be made and validated by human experts, today’s big data analysis is more “automatic” and involves less human involvement. The development of corporate big data solutions has allowed the rise of new platforms. Due to the ever increasing volume and complexity of data, today’s advanced corporate analytical workload very often includes the following four elements as been illustrated in figure 2:

- Streaming analytics
- Analysis of structured data
- Analysis of data not modelled and multi-structured data (i.e. Twitter for sentiment analytics)
- Graph analysis

![Figure 2 Main Elements in Corporate Data Analytic Tools](image2.png)

Both large corporates and small to medium enterprises use different types of data tool at various level of depth. Big Firms often have the financial capabilities to settle a large sized data analytics while small companies usually don’t have big data capability due to the extremely high-cost to design, customize build, and to improve the system. It very often involves the involvement of research institutes to help the companies to develop a system that could serve the purpose. Very few mid-size companies have the capabilities to develop big data solutions.
Companies may fall into the following four categories of big data practices (McAfee et al. 2012). 1. Big data competitors represent the organisations that have a full access from operations and are able to use the data available. 2. Overachievers have limited access to data but they try to use it as much as they can to make a decision. 3. Underachievers have full access to information but they don’t effectively use them to obtain strategic advantage. 4. Disadvantaged companies have limited access to data and no use of the data. Obviously, for companies ready to invest in big data solutions, their best outcome is to be a competitor.

We are interested to testify these observations, opinions, and models based on a study of the CAC 40 companies. The similarity in the CAC 40s is that they are all publicly listed large companies that cover a much diversified pool of industries. Because all CAC 40 should have the financial capability to invest in big data, we eliminate the large bias from the financial constraints. In summary, we want to provide evidence to prove the following two hypothesis regarding the depth of data involvement and the company’s long-term and short-term financial performance.

Hypothesis 1: Deeper the depth of data involvement, the better the short-term financial performance.

Hypothesis 2: Deeper the depth of data involvement, the better the long-term financial performance.

2.1. Measurements
To accurately quantify the depth of data involvement is not an easy task because of the lack of publicly available information of the implemented information system and its technical details. Implicitly, we therefore define the data depth of a company by examining its corporate structure in our first step. Based on the impact of big data implementation on corporations highlighted in the first part, we define big data commitment as the resources allocated to big data within the company. In current stage of our research, we focus on corporate governance factors. This organizational measurement is qualitative but it does offer a more accurate measurement than other quantifiable financial variables. Financial investment could have been a good indicator but most companies do not disclose precisely their investments in big data. Those investments take really diverse appearances (capital expenditure if they acquire data centers or software licenses, operating expenses if they use SaaS/IaaS as a support functions, COGS if they use SaaS/IaaS as operational tools, HR costs if they hire data scientists, operating costs if they hire external consultants, etc). Corporate governance is disclosed by all CAC 40 companies, allowing us to bench on homogeneous criteria their “big data commitment”.

Because in general the “big data commitment” takes place at the company-wide level, we therefore first study the corporate governance criteria and based on the following two factors: the presence of “Big Data Officer” and the presence of a “Big Data Chief Officer” as one of the Executive Committees of the company. We use 2 metrics to measure the financial performance of the CAC 40 companies: the stock price variation on the last one year and on the last 3 years.

We choose to use the general “stock price” as the relating financial performance indicator. We choose not to use pure profitability metrics such as Gross Margin, EBITDA, EBIT and Net Margin because they display more dynamics and inconsistency over time from different accounting procedures (of investment, asset depreciation and valuation, deferred income, etc). Another reason is that the high installing costs related to big data projects can cast significant impact on these financial metrics depending on the accounting practices and the way corporations organize their business units and headquarters reporting. Stock price also has the advantage to reflect future expectations. As we assume that big data projects will deliver the most of their impact in the near future, we consider stock prices in the past “one year” and “three years”.

It could also be possible to consider the stock price change “one year before” and “one year after” the initiation of the big data projects. We consider that the precise determination of the starting point of this period is nearly impossible, as an internal corporate initiative does not necessarily correspond to the time when they are made public. Additionally, some CAC40 “Big Data commitment period” took place very recently, preventing us from choosing a homogenous period for our data sample. Choosing different time periods (even if with the same length) would have resulted in heterogeneous comparison because of the noise generated by the impact of financial markets variations. As a result, the 1-year and 3-year time periods represent the short-term and medium-term indicators that correspond to a reasonable timeframe when CAC40 really began to consider Big Data as key tool for their business.

2.2. Findings
The chief data officer (CDTO) is considered one of the key executives of a company who manages the company-wide data processing and knowledge mining
and usually reports directly to the CEO. In this research, we consider those companies with a current CDTO position and those whose chief data officers have just left positions.

The chief digital officer (CDIO: in French the Vice-President en charge du Numérique or Vice-President en charge de la digitalisation) is another type of key executives of the company. CDIO mainly helps the company to digitize the old system by introducing online technologies, mobile applications, social media, electronic commerce, etc. Companies with a CTO with a clear implication in data projects are considered in this cluster. We also include in the cluster those companies whose chief digital officers have just left positions.

Although there are responsibility differences between the CDTO and the CDIO, usually there is only one position in a company. We use CDO to represent both CDTO and CDIO. This paper presents preliminary results of 1-year and 3-year financial performances. Because big data is a new phenomenon to most companies, we cannot find any useful measurement that can present long-term financial performance longer than the 3-year range.

Table 1 one year and three year stock performance Comparison

<table>
<thead>
<tr>
<th></th>
<th>CDTO</th>
<th>CDIO</th>
<th>NonCDO</th>
<th>CAC 40</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Year Stock Performance</td>
<td>-8.5%</td>
<td>-8.9%</td>
<td>-11.1%</td>
<td>-9.6%</td>
</tr>
<tr>
<td>3 Year Stock Performance</td>
<td>47.7%</td>
<td>36.9%</td>
<td>34.8%</td>
<td>37.8%</td>
</tr>
</tbody>
</table>

The analysis of both one-year and three-year stock performance show that companies with a CDO outperform the CAC 40 on average. In addition, those companies with CDTO perform slightly better than the ones with CDIO, but both of them perform better than the CAC 40 index, Table 1. The results are consistent in the past one year and three years.

Table 2 summarizes the stock performances of those companies with a CDO in the executive committee and those without. In both one year and three years, companies with a CDO at the executive committee outperform those companies without a CDO and outperform the CAC 40 index. For those companies with CDOs, the ones with CDO in the executive board perform better.

Overall, the results show average stock outperformance of the companies that are committing resources to big data implementation in the CAC40. When we examine case by case for each company, we also find some CDO companies that do not perform better than the index. We also discover some nonCDO companies perform much better than the index. We are more interested to look into these companies and give an explanation that may change the way we understand the role of big data in today’s corporation practices.

In general, our preliminary study shows that big data commitment in a company does create value that is reflected in the stock price on average. In what follows, we will focus on a more strategic approach to understand how big data creates value for a company and why some companies without CDO still perform well and some companies even with CDO perform less.

3. Big Data Strategies

Existing big data strategy model can be found in (Davenport et al. 2012), the strategic perspectives include: the access to a data streams and the big data capabilities of the company, Figure 3. This is very data driven, in the sense that the model considers only data’s perspective to outline the possible corporate data strategy. The drive from within the company is greatly missing in this model, for example the operational and supply chain dynamics that is closely relevant to the data strategy. It doesn’t explain the questions that we have previously raised. We incorporate supply chain dynamics into this strategy model by mixing data streaming capability and the data analytical capability as one broad perspective of as “data capability”.

Table 2 CDO in the Executive Committee

<table>
<thead>
<tr>
<th></th>
<th>Executive CDO</th>
<th>NonExecutive CDO</th>
<th>CAC 40</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Year Stock Performance</td>
<td>-4.8%</td>
<td>-10.7%</td>
<td>-9.6%</td>
</tr>
<tr>
<td>3 Year Stock Performance</td>
<td>48.1%</td>
<td>35.5%</td>
<td>37.8%</td>
</tr>
</tbody>
</table>
3.1. Value creation From Optimized Decision Making

In the value oriented big data strategy, we argue that companies still need to focus on the true value of the operations despite the huge “hypes” created by the big data. We must take into consideration both operational and supply chain uncertainty (dynamics) and the company’s data capability along with the company’s data capability (including both data acquisition and the analytical capabilities), Figure 4.

Based on the frequency and variety of decision makings in a company, we are able to categorize the operational dynamics as either low or high. We group data acquisition and analytical capability as a general term as “big data capability”, which can also be low or high. We thus create four strategies: routine, integration, strategic, and excellence.

Routine: This is the strategy when the operational activity in a company is stable and lack dynamics and if this company has limited capability to gain access or to analyse data of large volume. Many of today’s raw material suppliers, for an example, fall into this category. Data request is at a very low level and data processing is a routine as a result. Figure 5 shows the dynamics in the number of operational decisions of a data routine company, which is characterised by small band of uncertainty.

Integration: If a company has limited data capability while still facing a large number of operational dynamics and decision making requests, the data integration strategy is right for the company to handle the data need (Demirkan and Delen 2013). Public hospitals in France, for an example, are well known for the large number of demands and the delay of providing services. These hospitals are also known for the lack of financial resources to implement advanced IT systems to help improving the difficult operational situations. The integration strategy would suggest integrating the data capabilities from the suppliers, such as the equipment manufacturer, labs, blood bank, etc. Data integration could improve the stress that comes from data need and limited data capability.

Strategic: When the company has a high level big data capability in order to handle a large number of dynamic decisions, big data becomes strategic. For example, the fast fashion industry falls into his category. Zara has already strategically utilized its big data from all resources in order to create a fast responsive marketing campaign. With high level big data capability, almost all business processes and associated decisions can be made based on the
assistance of data analytics. For example in design process, if a company wants to decide the colour of a product, they can use the “A/B testing” to predict which colour would have the best impact on the market. Other data tools can be equally used to design the form of the product, to decide the right portfolio of the size, etc. Consequently, the big data is the strategic element.

Figure 7 shows that the dynamics in the operational decisions can be greatly reduced by the big data solutions due to the increased accuracy of prediction as well as the real time analysis and control.

Excellence: For companies with low operational and supply chain dynamics, big data solutions can be used to achieve “excellence” in service provision, quality management, customer satisfaction etc.

From the strategic transition perspective, we can use the model depict how companies can transit over time, Figure 9. For an instance, a routine company can adopt big data capability to achieve excellence, by using the data advantage to reach the best business and operational potential. For companies with high operational dynamics but with limited data capability, they can integrate data capability from suppliers and buyers to make it strategic or to reduce operational dynamics and makes the operations routine.

Companies with high operational uncertainty and with high level of data involvement can choose to use the data capacity to reduce the dynamics and achieve excellent operations. For example, plants that adopt a large scale of automations don’t necessarily need to make too many decisions that need human interferences.

3.2. Big Data Enabled Value Chain

In section 3.1 we present a new value oriented corporate data strategy schema. In his section, we highlight the know-how of value adding data strategy on the value chain and decision-making processes.

Figure 10 demonstrates the data integrated value chain. Real time data is created from inbound logistics, productions, outbound logistics, sales and services. The data is screened and processed according to the corporate strategy, in order to facilitate the main business components, including procurement, technology, human resource management, and organizational structure design. Through this process,
knowledge is created and is further utilized to support the daily operations and management.

Today’s big data system also involves real-time data acquisition, processing, and decision making. Real-time data allows decision-makers at the primary-activity level to benefit from real-time decision support. At the higher level, decision-makers benefit from strategic support. Those insights can improve the firm’s performance at each level of both primary and secondary activities.

At this point, we have examined the financial performance with CDO and without CDO and explained the role of value in the big data strategies. We now dig further to scrutinize the impact of big data and its strategic implications in “the real world” by a case study of airline industry’s big data practice.

4. Case Study: Big Data in Airline Industry

Big data creates value across a wide scope of industries, business models and competitive situations. We systematized our approach in order to make it reusable as a management tool by formalizing the gain from big data as “gain in efficiency” and “gain in knowledge” throughout the value chain. In the airline industry, newly developed tools based on data allows the managers to reduce the frequency of decision making, better understanding of the customer behavior, customize and optimize services.

By better anticipating various sources of potential problems and by reducing the risk of qs, big data decreases the frequency of decision making. Both managers and employees can focus more on the value adding activities. The airlines industry can be classified as “data strategic” with a large number of real time decision requirements as well as with a large selection of existing data tools. An airline company faces a large number of decisions: management of the flights (delay, technical problems, etc), management of the customers and the relationship, baggage tracking, food supply, airport coordination, etc. An effective data analytics would allow airlines to be more reactive and efficient in the management of their activities. Big Data increases the ability of the airlines to be very reactive in case of problems.

Figure 11 illustrates the current big data practices of airline companies (sources come from Lufthansa, Frontier, Delta, and AirFrance). Data comes from six major categories: airline data, frequent flyer, social media, aircraft data, users’ data and third party data. These data are effectively utilized to restructure the sales business model, to improve operations, to leverage the profit revenue optimization (PRO), to design better route plan, and to boost loyalty program.
4.1 Big data as a way to restructure sales business model

The airline industry is one of the most competitive industries. Nowadays, every airline is looking for opportunities to improve their revenues, their operations, and boost customer loyalty. Implementing big data solution in this industry becomes a leveraging tool. Currently, the airline market is divided in 3 main segments:

- The Full Service Carrier: “traditional” airlines such as American Airlines, Etihad, AirFrance-KLM.
- Low Cost Carrier: new competitors in this market. Their business model is based on the reduction of costs in order to offer a cheap price for the passengers.
- Charter: seasonal flight mainly linked with tour operator.

Many airline companies are seeking to explore new business models among the above three or something new. In general, the passengers are the key element of the business model selection and they are normally looking for the best price possible in the market. The electronic platforms allow passengers to easily switch from one company to another. With big data, airline companies are able to understand their passengers’ profile and behavior. Regarding consumers’ information, to be more specific, it includes the historic of payment as well as the willingness to pay, the time that it takes to make a purchase decision, the search path, preferences for the seating and food, geographical information, among others. Business models are supported by the management of operations.

4.2. Big data helps airlines improving their operations and management

The operational side of an airline company is very complex that involves many difficult problems (both managerial and mathematical). From security check, logistics, route planning, maintenance to refueling, automatic sensors and information communication technologies create a large volume of real time onsite data that can consequently help engineers and managers throughout the flight service. According to the International, Air Transport Association, the main reason for delay is airline-controlled processes (for instance maintenance). Consequently, if an airline is able to “see” all the conditions and make quick decisions, it would greatly increase the operational efficiency by keeping its fleet in the air. A plane costs a lot when it is on the ground (approximately 10K€ per hour based on the IATA analysis).

Big data could also help reducing unplanned downtime when an airline uses preventive process based on forecasting and pattern recognition. Collected historical data can help maintenance technicians execute the right work steps with the right tools. Currently, all the new aircrafts have a large number of modern sensors to record flight and mechanic data.

Delta Air Lines is one of the major airlines in the US. Delta’s success is somehow due to its strategic investment in big data technologies years ago. For instance, Delta has invested over $100 million just the in airport baggage systems to improve its baggage handling process. The operations team analyzes and understands the reasons of mishandled bags and integrates human expert knowledge with the automatically created data knowledge. As a result, Delta has developed a solution to know perfectly where all the bags are and where they should go. The system alerts baggage handlers when connecting bags that need to be directly transferred. Passengers can even follow the progress of their bags at the real time. Since 2007, Delta has reduced its mishandled baggage rate by 71%. This impressive improvement creates higher customer satisfaction and contributes to the increasing customer loyalty.

4.3. Big data is a leverage to optimize PRO

Almost all today’s airline company use the revenue management system to offer the best possible price. The competition is fierce and it seems extremely difficult to get a competitive advantage just by using this system. Several airlines have successfully adopted big data analytics to make a change. These data are effectively utilized to restructure the sales business model, to improve operations, to leverage the profit revenue optimization (PRO), to design better route plan, and to boost loyalty program.

Lufthansa is currently using a big data system to analyze and determine its price structure in real time rather than in the batch mode. The system allows Lufthansa to load more data and application functionality into memo and improve financial performance.

Frontier is a comparatively smaller carrier based in Denver. To survive and compete with larger carriers, Frontier developed a model by integrating all the tickets price of competitors and thanks to the analysis of those price, they were able to create niches of
pricing advantages for short periods (just before the competitors can respond).

4.3 Big Data helps creating ancillary revenues

In a highly competitive environment, the search for new service possibility (and consequently new revenues) is very important. Big data can be very helpful in this case. For instance, Southwest Airlines is using anonymous customer data to promote new products and services. New products and services come from deeply understanding the customers’ need, which creates real opportunity for the airline companies. Big data can facilitate to answer those small questions such as: do they need assistance before and after their travel, do they need a car, do they need a guide for their arrival, etc. All these questions can be answered from the analysis of big data, and much quicker.

Based on the analysis made in the retail industry, where constructive data analytics allows to increase sales, airlines used the same process to understand more easily their customers’ needs. Consequently, airline ancillary revenue can be developed to a large range product and not only baggage fees: the carriers can develop products based on real-time information about customer preferences.

4.4 Big data boosts loyalty program

All airline companies understand that to delight customers and to build loyalty is the key to success. Qantas is among the first to use big data to boost its loyalty program. Firstly, the Qantas airline has created a dedicated team that manages their social media channels and responds to customer complaints and posts in real-time. Consequently, Qantas in flight staff uses mobile devices to get access to customer profile and have a look to the historical data of each customer with the loyalty program. The staff is able to identify the top 10 customers for each class and they are able to give more attention to those specific customers.

Moreover, the capabilities of data analytics of Qantas are now well-known in the industry that they can even sell the system to another airline company. The mechanism is to get at least 10% of their frequent flyer members to authorize the carrier to share their data with third party clients. Qantas has already signed up clients like Hilton Hotels, Avis Car Rental, and American Express to expand its data network.

Figure 11 Big Data Schema in the Airline Industry

Figure 11 summarize the big data schema in the airline industry as what we have described in previous sections. Moreover, based on our proposed strategy, we can say that the airlines industries have moved from the “integration” strategy to the “strategic” strategy. Some airline companies even succeed in transiting to the “excellence” strategy. Using Big Data is a key element to develop the business activity of an airline and this industry has developed quickly its new technology. Nowadays, we can consider that this industry is one of the leaders. The different examples of Big Data use show that this technology can impact several aspects of your business and create value.

5. Conclusions

Big data is emerging in recent years as an evolutionary phenomenon. Many new concepts and business models driven by data are introduced as a result. We in this research are motivated to investigate the value side of big data. We examine the financial statements in CAC40 companies in order to discover the relationship between stock performance and depth of corporate data involvement. Our results are surprisingly two-folded. There are companies with strong data capability that succeed in the stock market. There are also companies without much data depth that also perform well. Moreover, the result doesn’t show link between the not-well performed companies and the lack of data capability. To decode this surprising result, we examine the existing big data strategy literature and discover the missing puzzle pieces. We thus propose a new strategic model that considers both supply chain
We find that big data can impact every part of business although the value that it creates greatly differs. Thus it becomes really important to understand the big picture of the business as well as to know the dynamics of the decision processes. Only by this means, companies are able to implement customized strategy that is tailor made for their specific value chains. There are some limits in our research that can be further improved as future research. First of all, we are not clear whether those companies invest resources in Big Data create a higher financial performance or whether companies with a higher financial performance invest in big data just because they can. Secondly, it’s also arguable that the share price evolution can be biased because shareholders are influenced by the “hype” around the big data project. We have to assume that shareholders are rationally pricing the companies with a “big data commitment”.

References
