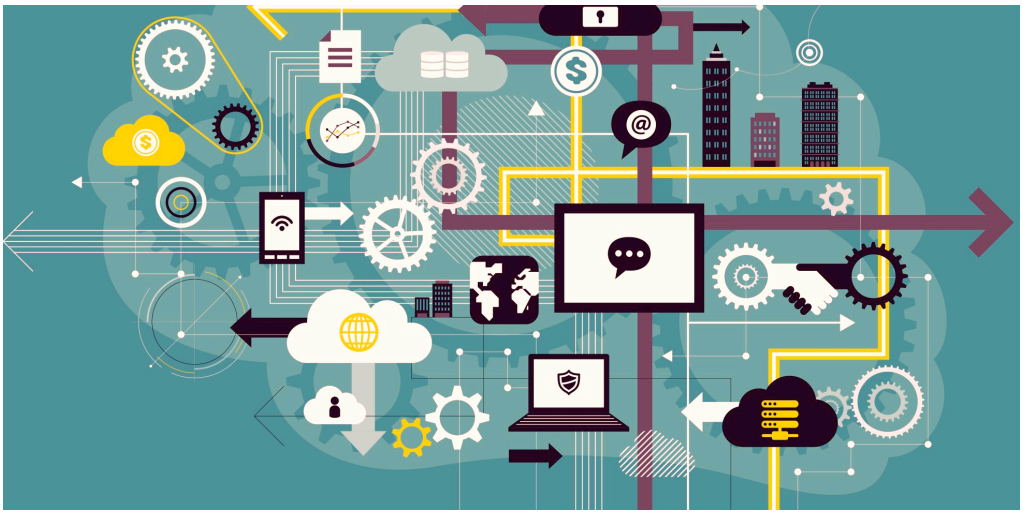
Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Economics

The Power of Big data

- How Big data affects organizational performance in Chinese dairy industry

XIA TIAN



Master's thesis · 30 hec · Advanced level
Environmental Economics and Management – Master's Programme
Degree thesis No. 1054 · ISSN 1401-4084
Uppsala 2016

The Power of Big data
- How can Big data affects organizational performance in Chinese dairy industry

XIA TIAN

Supervisor: Konstantinos Karantininis, Swedish University of Agricultural Sciences, Department of Economics

Examiner: Karin Hakelius, Swedish University of Agricultural Sciences, Department of Economics

Credits: 30 hec
Level: A2E
Course title: Independent Project in Business Administration
Course code: EX0782
Programme/Education: Environmental Economics and Management, Master's Programme
Faculty: Faculty of Natural Resources and Agricultural Sciences

Place of publication: Uppsala
Year of publication: 2016
Cover Picture: Datafloq
Name of Series: Degree project/SLU, Department of Economics
No: 1054
ISSN 1401-4084
Online publication: http://stud.epsilon.slu.se

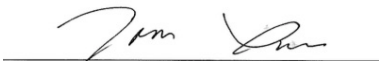
Keywords: *Big data, Business impacts, Digital technology, Information management*

SLU
Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences
Department of Economics

Acknowledgements

I would like to express my sincere gratitude to my supervisor Mr. Kostas Karantininis from Department of Economics at Swedish University of Agricultural Sciences. Thank you for all inspiration and constant support on my thesis process. You also provide quite effective guidance on this new topic area. I have learned lots of interesting dynamics from my thesis work and lastly, thank you for your encouragement in my most difficult time during each step of my thesis.

I would like to give a grateful appreciation to my cousin Mr. Du who has paid a great concern to my thesis and actively helped establish smooth connections with my interviewees. I also show my greatest thanks to all interviewees those of whom have taken so much time off your busy schedule. Finally, thanks to my dear friend Shizhi Wang, at Department of Animal Breeding and Genetics at SLU who takes so many concerns to discuss with me about the ideas, thoughts, and use of applications. Thank you for all these things you have spent to support, understand throughout my difficult process.



Abstract

The popularity of Big data system has increased in recent years due to the high speed and huge volume. Big data has changed firms the way to evaluate the marketplace, research customer demands and reconsider their business performance. Big data now is able to renovate the business model and reshape the structure in the market. In recent years, Big data has greatly influenced the organizational model in the Chinese dairy industry. But how the Chinese dairy firms grasp the Big data development opportunity and what results would it be?

The objective of the study is to identify with the current insight on which level the dairy companies in China are initiated with Big data in order to investigate on how Big data can affect the organizational performance. And also to investigate what are the advantages and challenges for using Big data as well as the how Big data influences the decision-making in reality.

For the theoretical part in this study, the focus is on organizational theory, the concept of digital technology, characteristics of Big data and business impacts of Big data. The qualitative research methods including literature review and in-depth interview have been used to know the situation of Big data implementation in Chinese dairy industry. The information management cycle has been used to the research in order to describe the empirical results.

The research analysis has discovered that Big data allows dairy firms to add value to customers and decision-making process. However, Big data will raise the risk of disclosing the corporate privacy information and high costs for establishing Big data system will place a burden for most of the small dairy firms.

Abbreviations

DT = Digital technology

TIR = Technological IT Resources

HI = Human IT Resources

IM = Information management

MIS = Management information systems

Table of Contents

List of figures and tables

.

1 Introduction

This chapter firstly introduces the current situation of Chinese dairy industry. The situation divides into primary producers and the supply chain. The first chapter briefly shows how known problems come from primary producers and supply chain that exist in the Chinese dairy industry. Next, the concept of Big data, which is the foundation of this study area, is presented, defined, and then visualize the problems where they occur in the dairy industry. The research problem is then described to clarify this study and the aim of this study is expressed; the following research questions are presented. Lastly, to help get a clearer overview of what this study concerns, the delimitations are easy to be presented and judged.

1.1 Problem background

In the past decade, driven by fast economic growth and drastic changes in civil dietary patterns, Chinese dairy industry has undergone a rapid increase. According to Wang (2000), dairy products increased from around 8.3 million tons in 2000 to about 38 million tons in 2008. In other words, there is above fourfold amount of dairy increase in the last fifteen years, and nowadays China becomes the third largest dairy producer in the worldwide range, just after India and the States. Compared with the quantity of dairy, the number of cows only grows with 2.5 times from 4.9 to 12.2 million (*ibid*). The quick growth of dairy production mainly arises from the expansion of growing number of cows in the first hand; and secondly, the technological progress. Specifically, the primary dairy producers in China have adopted more high-yield breeds, like Holstein cows, which are generally adopted by the dairy farmers having more capital, productive knowledge and large-scale breeding base to handle complicated production procedures for high-yield breeds (*ibid*).

Previous research (Pehrsson *et. al.*, 1999) finds that the majority of primary dairy producers in China have gained the prime time of development in the last century. Some of the pioneers such as Yili, Mengniu, and Guangming have gained vigorous development from the national support and local investment. Furthermore, the study shows that there is a considerable achievement in both production and sales fields from the primary producers. As the statistics revealed by the Chinese Dairy Association (2008), there is about 67.5 percent cows in China are raised by small-scale dairy farms and these farms normally own less than 20 cows in one farm. On the on hand, even if the number of backyard dairy farms, which only raise less than four cows per farm, occupies 81 percent of total Chinese dairy farms, can contribute about 40 percent of total national dairy output (*ibid*). On the other hand, the number of large-scale farms (more than 100 cows per farm) takes only 0.3 percent of total dairy farms, but those farms can produce nearly 17 percent of total dairy output in surprise (*ibid*).

The booming scenario of the Chinese dairy industry results in the improvement of supply chain since hundreds of the dairy companies have to develop it in order to procure the sales volume. The dairy supply chain in China is experiencing much more rapid revolution than anywhere in the world (Hu *et. al.*, 2004). According to their study, both small and large farms have equal opportunities to take part in the modern marketing activities in China. Lohmar *et. al.* (2008) identifies that the majority of dairy companies in China have modern processing plants with internationally-recognized quality management systems in place, but dairy companies have to rely on supply chains which embrace a network between producers and traders, rural supply stations, and approximately 2 million dairy farmers in the remote and

under-developed regions. In China's northern provinces, the areas have abundant land occupied by small-scale farms where they eager to supply milk, which can help yield a stable monthly income (*ibid*). The large-scale farms have to confront with fierce competition at all stages of the supply chain, and the market dynamics at different levels are at times out of sync (*ibid*). The excessive expansion in Chinese dairy industry leads to the continuous growth of milk price in supply despite of the surging demand. However, a series of problems occur such as over capacity, technical upgrade difficulty, and low efficiency are arising with the over-expansion in Chinese dairy industry.

It is obvious that companies need to capture a huge amount of information consistently. But with the popularization of the Internet, along with the efficient use of computational tools and digital storage, the amounts of data have exploded in recent years (Science, 2011). IBM has disclosed that we can create 2.5 quintillion bytes of data every day from the mouse click and finger swipe (*ibid*). Data, regardless of its category, location, and origin, has become a core business treasure for most of the enterprises since it can directly reflect the business scenario from both internal and external to the business environment. As a result of the increased awareness in companies and customers, the concept of data has been developed along with the time progress. With other words, it is no doubt that Big data is applicable in the statistical fields such as calculation and communication. The use of Big data's statistical pattern can predict behaviors and processes in the large variety sectors of business. And in real cases, Big data has been shown its value and usability when it comes to productivity, innovativeness, and profits to the company. Nevertheless, it is predicted that the massive use of Big data may bring about problems which relate to information privacy, personal safety, and utility. Thus, many people argued that overconfidence in Big data may become more risky in societal aspect (Nordström, 2014).

1.2 Problem

From a theoretical perspective, there has a controversial debate on what the impact of technology data in company's success is and in which way the data makes the contribution to company's effectiveness and performance, but the result seems rather unclear (Mata *et al.*, 1995). Even though the awareness of the concern of company's performance, scholars and academics do not seem to consistently agree upon which the interaction between data and company performance is, and especially to how data be regarded as the catalyst to company's achievement (Nevo, 2010). On the one hand, it was suspected that technology data are even unlikely to be deemed as a crucial factor to company's success since data and information are transparent and timeliness, which they are widely available to all of the users in the competitive market (Mata *et al.*, 1995). Nevertheless, on the other hand, it was realized that technology data has been demonstrated to offer abilities to which, when associated with other factors in the company such as human resource and transaction knowledge, can give a basis for recognition of an organizational performance (Kros *et al.*, 2011). As mentioned by Nevo (2010), data has become the tangible asset to companies in recent years, it is not fundamental to which companies are adopting, but how they implement data into their real activities. This ensures the thesis approach to stress on how Big data are used and assume to make the organizational performance more active, because the assumption affects the way in which the organization uses Big data in general.

To consider the fact that Big data is a more specific example of technology data, and the debate in recent years seems to be more drastic when Big data is regarded as the powerful tool that it is able to produce large amount of data which is all-dimensional within the company

(Yue *et al*., 2014). Companies can access to most of the information without difficulties so far; however, the path between Big data and the assessment of performance in organizations is rather unexplored. The vision of this study is to investigate firstly to what extend Big data is used in Chinese dairy industry and next, how Big data are becoming applicable and perceived to affect the company's decision such as the strategic plan, management adjustment, *etc*. By using the theory developed so far about Big data as instruction, the study will explore a quite recent and new study area, the Chinese dairy industry, as one of the study object for Big data. The study will center on the dairy industry in China, because it is a fully attractive area with high level of interest due to a big amount of data generating for future analysis (Liang, 2014). As mentioned before, once Big data is adopted by companies from various sources like customer's data, it would be considerably effective for companies to get insight into purchasing behaviors, personalize offers, and increase the margins (Savitz, 2012).

Knowledge on how the effect of Big data are identified to lead to organizational performance that will result in deeper insight into the existing theoretical problem about how data can bring benefit to company's real performance. Since the problem concerning the contribution of data to company's effectiveness and performance is still vague (Mata *et al.*, 1995), empirics on how organizations consider this new and powerful Big data, will also lead to the existing argument. The study also involves with information management by promoting theoretical and practical understanding of how Big data affects the organizational performance in Chinese dairy industry and it also helps to understand the whether Big data is applicable to a company. The empirical implications, which are valuable for companies taking part in the usage of Big data as it provides a great overview of the usage phenomenon so far, and helping companies on what to expect about the use of Big data and how this will influence their performance. The study will provide suggestions to those dairy companies interested in Big data and to have an awareness of changing paradigms on how those can consolidate their organizational performance either before investing in such technology, or also after, in understanding how it should be used in the appropriate way in order to lead to a better organizational performance.

1.3 Aim

The aim of this study is to identify with the current insight on which level the dairy companies in China are initiated with Big data in order to investigate on how Big data can affect the organizational performance in Chinese dairy industry. To reach this aim, two research questions are presented.

-*How, and to what extent, do Chinese dairy firms apply Big data in their business?*
-*What are the benefits or hindrances of using Big data in the dairy organization?*

1.4 Delimitations

In order to conduct this study, it is necessary to specify the delimitations. Firstly, this study has been delimted to one specific industry field, the Chinese dairy industry. There are three different established dairy companies in China have been identified; in unfortunate, one of these dairy companies was not able to contribute positive information of using Big data in this study. The rest of two dairy companies (Mengniu and Yili Dairy) have contributed with more active information about their situation. Two interviews have also been done with the company manager or CEO of conducting Big data in order to receive more information about

the Big data situation in Chinese dairy industry. The focus of this study is to identify the usage of Big data and how it can affect the organizational performance in general which accordingly provides valuable empirical information and deep understanding about the enabling factors for Big data.

This research is carried out from the organization's perspective. It means that the focal point of the study, both interviews and the literature review are typically based on the dairy companies in China and the company's perspective. There is no empirical data coming from dairy companies located on the outside of China since they are not a part of our focus group. Nevertheless, it can be noticed that most of the existing literature related to Big data and company performance comes from outside China, which gives suggestions of the Big data situations in other developing countries. Data derived from the three contributing companies have been collected, and the situations of the companies are different from each other. Since the usage of Big data in Chinese organization is very little, a collection of data from the case companies seemed suitable and valuable. The informants interviewed have been completed in person, only by telephone channel. One main interview has been split into two different people in Yiyi Dairy, but in that case, additional questions would be asked to clarify specifics.

Delimitation regarding theoretical perspectives is also distinct. Since the study applies to use the explorative and inductive approach which leans upon developing potential theories about underlying structure that are transparent in text data. The thesis delimits on relying observation for information collection thus the universe scenario of Big data in the dairy industry needs to have an in-depth analysis from the theoretical divisions.

2 Literature Review and Theoretical Framework

In this chapter, the concept of organizational theory, the theory regarding digital technology, information management as a business model and the theory of the Big data, and its impact on business performance are presented. The reason for presenting these and to illustrate a conceptual framework is to increase the scientific understanding of the wide concept of Big data. The theories and concepts presented here are later discussed and analyzed along with the empirical finding to give this study an academic degree of depth.

2.1 Organizational Theory

The study is about the usage of Big data in business and identifies how dairy companies uses to improve their organizational performance. To learn how the company knows technology power, we have to know something about organizations and theory behind it. Organizational theories derive from organizational practices and in turn, serve the practice. According to Nicholson (1995), the organizational theory was defined as a series of academic viewpoints, which attempt to explain the municipalities of organizational structure and operating process. In another word, organizational theories are knowledge systems, which aim at finding organizational structure, function and operation, organizational group structure and individual performance (*ibid*). Daft (2007) elaborates that applying organization theory has advantages for both organization and society at large. And organization exist when people have interaction with one another to perform essential, they are social units of people with noticeable a boundary to meet with certain goals (*ibid*). The organizational society at present is totally different from which we lived in the old society focusing on agriculture production; information is isolated, closed, and individualized (*ibid*). We have then gone forward to an industrialized and information society using shared knowledge to serve oriented organizations, where the way we act and communicate with other people becomes more important that the real product from organizations (*ibid*).

In the early of last century, the classical organizational theory came out with the development of industry dynamic (Guo, 2003). Taylor (1911) developed the skeleton of classical organization theory, by identifying such objects: individual efficiency, enterprise organizational efficiency, and social organizational efficiency. It was encouraged that the bottom-up rational administrant method, which can influence the people's working status by altering the way to carry out the individual missions replaces the arbitrary managers with scientific and rational procedures (*ibid*). In comparison with Weber's (1964) administration theory, his arguments think that authorities are the key concept in the administrative system. Weber considered the legal authority plays the essential role in the system in order to keep the enduring bureaucratic structure, which is more convenient to deal with the complex affairs than the traditional authority, because the the structure in legal authority has absolute technology superiority than the traditional one (*ibid*). In classical organizational theories, the organizational structure is deemed as the essential intermediary to achieve the bounded rationality, and it aims to attain the rationalization of organizational structure (Liu, 2010). Hence, the classical organizational theory is the common philosophy on human-to-machine perspective. In this background, the metaphors of an organization and individuals semms like gears in a machine and people are systematism, obedient, and lost their self-initiative at work (*ibid*).

The modern organizational theory has risen with the third technical revolution, and it brought many new social and economic phenomena and influenced organizational environments thoroughly (Liu, 2010). The knowledge of classical organizational theory is not able to explain such organizational changes anymore, and it requires the transformation of theories and brings into a new growing phase. From the human interaction perspective, modern organization theories treat organizations as open structure. It underlines the impact on the organizational environments and brings about new insights and methods from complexity science, and different research paradigm (*ibid*).

The organizational environment in a broad sense means the big environment including outside and inside which would potentially affect the organizational operation and performance (Duncan, 1972). For a specialized organization, the specific organizational environment is more important than general environment because the specific environment can directly affect the decision and action of the manager and is straight relevant to the accomplishment of organizational goals. However, the general environment is not usually been directly taken into account and often been treated as factors into management decision (*ibid*). According to Scott (2003) shows that organization and their corporate environment is correlative dependent. It has been considered that the organizational structure is the common feature of organizational operation and environmental reaction (*ibid*). In addition, the organizational environment restricts the organizational behavior, and organizations attempt to use the chances of institutions to enhance their business performance (*ibid*).

The study of the organizational complexity is mainly related to the complex science and it has grown to be a broad-ranging subject after it emerges (Xingtang *et al.*, 2008). For most reasons, managers feel pressure on new ideas and methods and management consultants are easier to put untested organization science thoughts into instant practice, which are susceptible to managers and management practice (*ibid*). In most time, the complexity science is more applicable in the areas of physics and life science since it can be seen as the organizational environment, all above will improve the applications of complexity science into organizational theories (*ibid*).

2.2 Digital technology theory on business

Digital technology (DT) is considered as the fundamental technique in modern organizations, and it has changed many ways of doing business. Morabito (2013) refers that "*The business value in DT suggests to the organizational performance impacts of information technology, including productivity enhancement, profitability improvement, cost reduction, inventory reduction, and another measure of performance*". DT also contributes to enhancing the organizational competitive advantage, which implies a distinctive market positioning that allows organizations to benefit from the above-normal profit, compared to the company's competitors.

To have the general understanding of how DT has the improvement on organizational performance, Morabito (*ibid*) has listed two major effects on organizational performance, which are the operational performance and strategic performance. Melville *et al.* (2004) have identified that DT resources embrace both Technological IT Resources (TIR) and the Human IT Resources (HI). Both TIR and HIR are universally applied to the business process, and it will lead to the organizational performance improvement. Recent studies of DT have examined how information technology can improve the operational performance of organizations, and they have suggested that DT resources can improve the business

performance from computerized reservation systems and also the relevance networks. Moreover, DT resources allow to having the improvements, which can be typically reflected on company's efficiency and effectiveness (Brynjolfsson and Hitt, 1996). These improvements are originated from effective of restructuring process architecture, better cooperation among disconnected departments and offices, identification of faults in the chain of corporate activities, minimization of errors through automatic processing, quickly observation on mistakes, *etc.* (Morabito, 2013). Organizations may obtain substantial improvements in their operational performance as the investment in DT will lead to cost efficiency and having good quality products, which in turn can perfect the company's entire financial performance.

The relationship between DT and strategic performance is still under a closer observation among different researchers (Luftman and Kempaiah, 2007). And some empirical cases reveal how the DT resource may improve the strategic performance in organizations (Melville *et al.*, 2004). One approach is to evaluate the impact of DT on strategic performance. And it entails measuring to which level of DT and organizational performance and DT systems are closely related. For example, in stock markets, companies react actively to announcements that firms are using strategic DT systems. More importantly, in the scenario of following such announcements, these firms appear more productive and profitable than their competitors (*ibid*). Other studies have suggested that those firms making input on DT systems may achieve more competitive advantage; furthermore, their established technology foundation represents a significant source of sustainability (Barney, 1991). In detail, achieving competitive advantage is characterized as the most desirable objective for a real company. From Morabito (2003), the general fulfillment of a competitive advantage requires being disentangled into its specific components. For a company on which focuses manufacturing and processing, the strategic performance of this company may be split into three different strategic targets: (a) cost reduction, (b) quality improvement, and (c) revenue-growth. Therefore, DT system provides a substantive assistance to each of those strategic targets. The empirical research conducted by Melville *et al.* (2004) shows that DT can support product and service differentiation or the innovation rate in order to promote the company's overall competitiveness. In short, DT resources may have a crucial organizational impact on the complementary firm performance such as practices, innovation, and structure. Hence, to be implicitly said, DT sources play a major role in shaping and sustain a company's competitive advantage.

2.3 Information management theory

DT once been created, has vested with its mission of delivering the right technology solution to their clients (Morabito, 2013). There are tons of thousands of books, journals, web pages, and television channels covering and promoting the future technological innovation in the DT fields. IT companies invest millions and billions of connections with their technological devices in order to ensure that their networks reach to every user's desktop. Perhaps among different uses, one of the essential issues is overlooked in DT industry – information. However, due to technological information improvement and the wide dissemination of information, many companies have to confront with information overload. Hence, information management (IM) is introduced to solve the information chaos in this digital world. In most cases, the information is stored in computer software in an unstructured form, spread in databases, rendering access to relevant knowledge difficult, and accommodate to employee's daily productive activities (*ibid*).

Barua *et. al*. (1995) have measured that most companies attempted to drift towards the form of information which is highly structured and easy to get access with. Besides, companies tend to collect automated data based on the aspects of their operational activities and new enterprise systems are mainly effective at processing and gathering structured transaction data. The data that been transformed into more structured form requires to have considerable intellectual concentration but most companies only have a biased view on its technological provisions (*ibid*). The defined terminologies such as data, information, knowledge, and intelligence can also bring about legitimate skepticism. For instance, those terminologies can be circularly used and having a mutual replacement and as if they are interchangeable (Morabito, 2013). Figure 1 below indicates the possible association between data, information, knowledge, and intelligence.

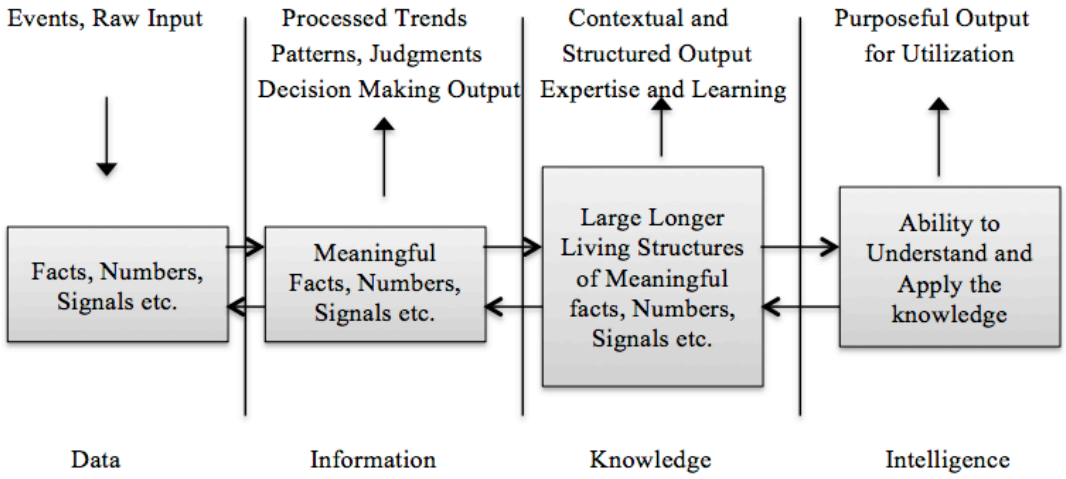


Figure 1. The association between data, information, knowledge and intelligence. (Morabito, 2013).

The detailed figure represents a panorama ranging from single facts, numbers, signals about human events and actions to more structured and meaningful concepts. Theses concepts include facts, trends, purpose or human perception, and contextualization of data in a given context of information, to concepts, which involved real reasoning. This reasoning generates new information, which based upon a set of beliefs that are at the same time been considered as individual and collective knowledge (*ibid*). In addition, the application of knowledge presented to reach cautious, sagacious decisions in the conflict scenario or the ability to learn and to apply knowledge.

Dias (2001) sets up a definition of IM as a transformation procedure of data into information, which is operated by the decision hierarchy in an organization. Management Information Systems (MIS) can provide more consistent information for the decision-making. At the same time, MIS is able to produce a joint view connecting present and past, and helping top managers to elaborate more realistic prospects. There are huge benefits where MIS can bring about from different aspects in the company such as improve productivity and quality, reduce operational costs, allocate decision-making process, and facilitate information access (*ibid*). According to Cronin and Davenport (1991), IM can also be considered as the conscious procedure by which information is gathered and use to assist in decision-making at all levels of organizations. From this definition, it can be found that there are three obvious characteristics regarding IM (*ibid*). First and foremost, IM is a conscious and consistent process and it has to be envisaged or performed in advance. This implies that IM has to be

designed, organized, and structured. The following characteristic shows that the objective of IM is to provide assistance in a decision-making process. As Cronin and Davenport (1991) claimed that information is not obtained for its own sake, but it needs to be gathered for the purpose of being used. After that, IM suits in the scenarios when the conscious planning process not with information but with the decisions have to be made. The third characteristic lies on the point that IM is for the benefit at all organizational levels. In enterprises, IM has considered as being a control mechanism, which can generate profits for managers or shareholders. IM should be at least inclined to add decision-making across and between all the organizational levels as it is for senior planners and decision makers.

To determine the effect of IM, Butcher and Rowley (1998) proposed the '7R's model of Information Management' which illustrated on Figure 2. This model shows an integrated information management cycle, from information reading to recognition, re-interpretation, reviewing, release, restructuring, and ultimately, to information retrieval. This cycle can be considered as a typical paradigm in any information environment, which also includes the digital world.

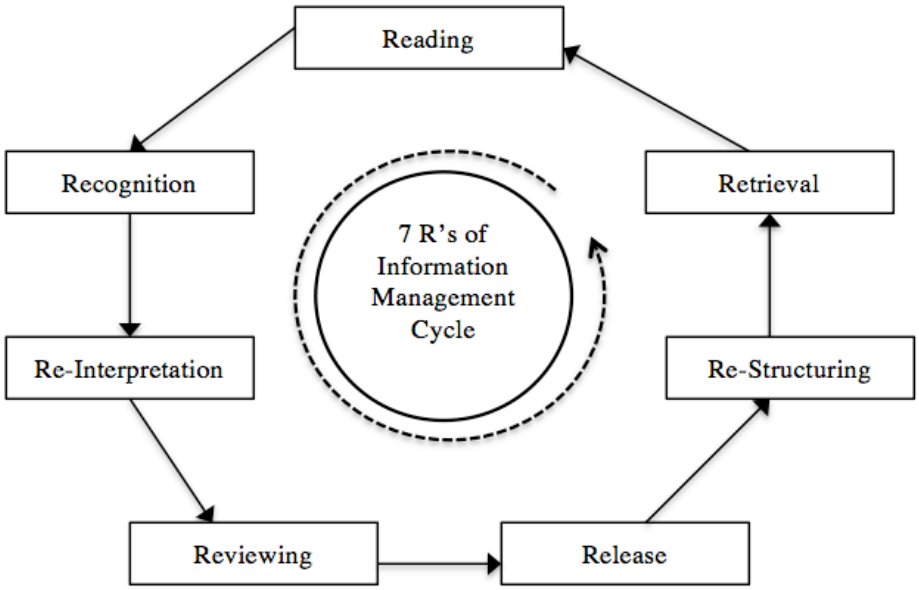


Figure 2. Information Management Cycle. (Butcher and Rowley, 1998).

The first step of information management cycle (*Reading*), as indicated from the figure, takes place when people read and obtain relevant data which recorded automatically in e-mails, documents, reports, etc. Such data is visually displayed on the portal web surface on the computer screen (Butcher and Rowley, 1998). As soon as we have read such information, the data will be absorbed into the cognitive frame in our mind. Next, information will be converted to subjective knowledge, and it will conduct a selective match between information content and user's individual concepts during the cognitive process (*Recognition*). The following phase (*Re-Interpretation*) occurs when subjective knowledge is transmitted to another vehicle. Moreover, this transmission can be stimulated by word processors, spreadsheets, software, etc. and retrieved by the collective processing component of the portal (*ibid*). (*Reviewing*) in this management cycle, is interpreted as the validation or assessment on what was transmitted by the individual user, or perhaps it is supported by office computerization and group software (*ibid*). Group users are facilitated to propose changes,

correct mistakes, build up links to other users that can share and deal with the similar subject, etc. Once the group users confirm the validation of information knowledge, it will then comes to public domain during the (*Release*) stage, that is, the knowledge becomes broadly accessible to any individual users in the community. The release of knowledge which occurs inside the enterprise may spread through the communicational network such as e-mails, virtual journals, bulletin news, *etc*. On top of this, the *Release* stage will take place through the specific web interface available on Internet when using an entry from the corporation (*ibid*). In most of the organizations, it is common to develop and store general information which is created by enterprise employees in operational databases, using systems, applications, and transactions from distinctive aspects. For managing and decision-making, however, that information which been measured having relevant and strategic functions will be extracted from these operational databases and then loaded into the information processing system. In other words, the (*Restructuring*) phase reflects the processes which comprise information extraction, transformation, and loading into a data warehouse (Morabito, 2013). Lastly, the *Retrieval* of relevant knowledge relies on information collection or repository, and it conducts with some retrieval techniques such as customized tools, tools that use metadata, business intelligence, and some analytical tools that compose the portal decision processing system.

2.4 Big data as certain characteristics

One article " A Brief Introduction on Big data 5V's Characteristics" created by Ishwarappa and Anuradha (2015), defines four different characteristics of Big data, which consist of volume, velocity, variety, value, and veracity. Because the characteristic regarding veracity has the meaning for redundant and junk information, which has little application in real practice; and some arguments concerning Big data characteristic request to omit the veracity element (*ibid*). Therefore, we had a minor adjustment to turn the 5V's model into 4V's in order to have a better comprehension in analysis chapter. The 4V's have the distinct traits respectively and these characteristics result in various paradigms and from different data lifecycle process is correspondingly try to achieve needed efficiencies. Figure 3 below will show the 4V's model concerning Big data's characteristics which are reconstructed by the author.

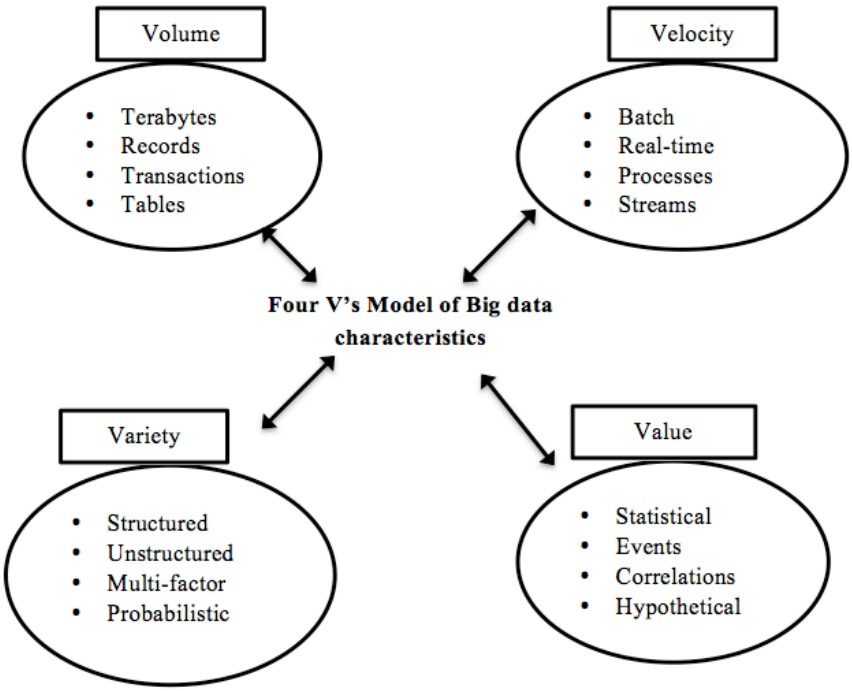


Figure 3. The 4 V's model of Big data characteristic. (Ishwarappa, P. and Anuradha J.; modified by the author, 2015)

- *Volume*

The volume of Big data refers to the quantity of generated and stored data that is being manipulated and analyzed for the purpose of reaching the desired results (Hurwitz, *et. al.*, 2013). It is one of the clear aspects that come to people's minds when refers to Big data. It also means a potential challenge because to manipulate and analyze a vast number of data requires considerable resources, which will consequently materialize in showing the desired results. For example, the processing speed can significantly limit the current computer operational system. However, the size of the data that is being processed can be unlimited, and the processing speed is unconstrained (Ishwarappa and Anuradha, 2015). The benefit that gained from the ability to process a large amount of information is the main attraction of Big data analytics.

- *Velocity*

Ishwarappa and Anuradha (2015) carried out their investigation on the Big data velocity. Velocity means the speed at which new data is created and the speed at which data circulate around. In other words, point A can be known as an end user interface or a server; data can travel from point A to point B, where B can have the same characteristic as point A, having its speed, is the form of data velocity. In the current era, data are generated in real-time or close the certain time frame. With the widespread use of Internet devices and wireless techniques, most of the machines and digital devices can pass on their data once they have started to work. Data flow rates are surged with massive speeds and variability, also creating a new challenge to enable the usage of real-time data.

- *Variety*

The following characteristics of big data are its variety. It describes the type of data that is stored, analyzed and used (Alexandru, 2013). Besides, big data is not always structured and also not easy to put such data into a relational database. In previous, the data variety was

disposed through transforms or pre-analytics to derive the results that would make the incorporation of other data in a relational model. To consider the fact that data formats, range, structures, and timescales are expected to use in analytics, the integration of this data would become more complicated. Nowadays, 90% of the generated data is unstructured, and new data appears in different formats: structured data, unstructured data, semi-structured data, and even more complex structured data (*ibid*). The extensive range of big data requires different approaches as well as advanced techniques to store all raw data.

- *Value*

The deep study concerning Big data characteristic by Ishwarappa and Anuradha (2015) argues that value is the most important aspect of the Big data. With right analytics, Big data can provide a tremendous insight and value to production performance, customers, supply chain, *etc.* Nevertheless, to harness such value in Big data, we need a fast and cost-effective solution that can transform Big data into accessible information (Oracle, 2013). For instance, it becomes costly to implement IT infrastructure systems to store Big data, and business will require a constant return on such investment. Additionally, these data may provide very little earnings if we are unfamiliar with its process. The fundamental aspect is to be able to turn this data into information that can serve to predict future and correspondingly makes the right decisions to cope with variation and risk.

2.5 Big data in different organizations

Information, where comes from the unconventional data sources, is attracted to the corporate environment since it is useful to analyze and extract the deep awareness from a maze of data set. Banica and Hagiu (2015) have introduced several patterns by means of which type of companies can improve the business performance by using Big data.

- *Great software companies*: software magnates such as Google, Facebook, E-bay, and LinkedIn, they have expressed the strong willingness to process Big data in their Cloud environmental several years before. They currently rely on their enormous integration capacity to collect information, analyze conventional measures such as sales volume on social media, as well as location information on their mobile apps.

- *Product development for Internet companies*: Big data is able to support such companies to capture consumer preferences and then, the reflected information will be proposed for designing and adjusting new products. For instance, Ford Motor *Ltd*. has launched the identical set of components, which would be assembled in Ford cars and trucks by implementing data algorithm that can consolidate more than ten thousand relevant comments.

- *Human resources*: Big data has the convenience advantage of handling the health care of employees in most of the companies. For example, Caesars Entertainment Corp. seeks and analyzes the medical-cure data for this company's 65000 employees, and Big data has found the optimal solutions on how to allocate medical services, the number of emergency rooms and whether they choose a generic clinic or brand-name drug.

- *Marketing:* Big data analytics provides a considerable effectiveness on information filtering and therefore the forecast of results shows more accurate and reliable. For instance, the Intercity Hotel SHG in Germany has collected details from 7.1 million guests

and members from their age, demographic traits, and income levels thus to predict whether they prefer family-style or business accommodations.

- *Manufacturing:* Superior manufacturing quality, as well as reduced production cost, is crucial to driving profitability to any major manufacturer. Big data can enable the combined analysis from angles of suppliers, customers, which promotes the analysis of yield and data quality (ORACLE, 2015). Moreover, data relate to throughput, material effectiveness, production attainment can be manipulated for future data analytics to seek for quality improvement together with cost reduction.

2.6 Big data on business impacts

As noted by Halevy *et al.* (2009), one of the most significant impacts of Big data is based on the organizational change or transformation necessary to encourage and explore the Big data opportunity. Meanwhile, Big data can reveal the potential benefits to any types of organization in any locations across the globe. Big data is not just superficially deemed as quantity or volume, but combining different data sets will provide organizations with real insights that can be used in decision-making and to develop the financial position in organizations (Katal *et al.*, 2013). In addition, organizations are using Big data to resolve the unanswered problems in order to refine and update their business performance and decision-making. Stephanie and Barbara (2015) from MIT have proposed an explanatory point for the improvement and innovation effects on which Big data created in the organization. The following context will emphasize how Big data has made an impact on the organizational model.

2.6.1 Improve the organizational model

As mentioned before, Big data has brought about an enormous potential value for solving dimensional problems in order to refine and optimize the business process and decision-making step. In this scenario, Big data supports the strategic measurement in novel ways by offering new data, insight, and action (Stephanie and Barbara, 2015).

- *New data*: it can be observed that manufacturers have chronically lacked the visibility of end-consumers. In previous, manufacturers have to depend on trade partners and data aggregators to release these data into their internal product development and customer-oriented progress. At present, manufactures can easily leverage Big data with the help of embed sensors for tracking actual product usage. Besides, they also make good use of social media in order to collect customer-produced feedback. Armed with new data and technology, these companies can easily generate business insights and provide with valuable information on customer preferences.

- *New insight*: Wixom and Ross (2012) have verified that there are lots of agencies decide to use Big data analytics for the purpose of identifying outliers according to data variation. In the experimental period, analytics can predict, estimate risk metrics for 45000 found observations (*ibid*). Furthermore, these results will be transmitted through visualization software to get a first-hand data sources. Consequentially, the organizational success is not solely based on Big data but it stems from new Big data techniques which have developed high-end statistics and models to diversify the insight of output.

- *New actions*: When organizations become well equipped with Big data and proficient on exploring resources based on that data, absolutely they will perform differently – be more intelligent and wise. The research conducted by Stephanie and Barbara (2015) has identified the new action as a result of employees working brilliant across the organization. An example comes from 7-Eleven, a company from Japan, who has invested on Big data with detailed sales and customer records, and put these metrics to its clerks using intuitive dashboard technology. This data has given a good opportunity to its customer's preference and generate a retailing miracle that 70% of all products sold are new products. The new action has promoted this company into one of the most profitable retailers in Japan.

From the study refers above; Big data enables improvements in business models across different types of industries. Big data can initiate a functional and strong strategic metric from the commodity based-practice. The most effective improvement originated from creating well -designed strategies that are informed by data and then honed and carved accordingly (Wixom and Ross, 2012).

2.6.2 Innovate the organizational model

The diverse functions of Big data, which can be used in organizations, are finding different ways of generating profits. Stephanie and Barbara (2015) has proposed two contemporary patterns which comprise data monetization and digital transformation. Besides, digital transformation happens when companies manage digitization to move into new industry fields or to create new ones.

- *Data monetization*: data monetization was interpreted as the "*act of exchanging information-based products and services for something of perceived equivalent value*" (Wixom, 2014). Products and services that added with informational data would enhance their attributes in complexity; capabilities and skills requirements; and the potential paybacks. It can occur in any stages at any level in processes such as raw data acquisition, data improvement, reporting and analytics, process design and execution. The evidence shows that in a digital company comScore which has achieved big success by helping advisers and marketers in organizations to engage digitally with the market segments (Wixom *et al.*, 2013). What is more, comSore has provided syndicated data products to its customers such as benchmark and demographic reports. The company's advice is based on how to make a quick response to insights on the background of marketing campaigns. Regardless the company character, data monetization also requests an organizational group who devote to build up the competence environment for the scope of its monetization business.

- *Digital transformation:* Digitization is not restricted to generate and increase the data volume. Also, according to Weill and Woerner (2013), digital transformation is an outcome of the business and organization which seek to adapt to this onslaught of disruptive technologies altering the customer and employee's behavior. On this basis, digitalization can be maintained as the development of organizational ecosystems, coordinated network of partner companies and intangible tools having added value to enterprises. Wixom and Ross (2012) clarified this term from a supplementary view that Big data can propel companies into a new field of industry or ecosystem, and it affects the traditional competitive landscapes. An illustrational example from Coles, the largest retailer company in Australia, introduced insurance services as it spread in dimensional offerings and became the leading sellers at present. Before the use of Big data, there are few insurers included Coles in their competitive analysis. However, Coles has currently

conducted a hand-to-hand combat with such enterprises. It suggests that traditional industry offerings are not the decisive elements in today's competitional environment. Those companies which can understand the customers best – likely from using Big data – will transfer a winning situation now.

From the study referring to Big data innovative function, Big data provides rich, exciting chances to leverage and extend an organization's strategic toolbox. Compare to most of the new technologies, Big data itself do not make strategic patterns. However, generating corporate value from Big data becomes more reliant upon the strategy makers and surrounding structures and processes in today's environment (Stephanie and Barbara, 2015).

2.7 Theoretical Framework

Nicholson (1995) describes that organizations are universal phenomena in human social and organization theories are derived from organizational practice and which in turn would serve practices. He identifies that organizational theory attempts to explain the municipalities of organizational structure and operating process. Aslo this theory is considered as knowledge systems, which aim at finding organizational structure, function and operation, organizational group structure and individual performance (*ibid*). The development of productivity and new technologies has resulted in the radical change of organizational environments and organizational practice that propels and perfects the organizational theory (Daft, 2007). In organizational theories, the classical organizational theory is firstly presented with the background of the human-machine relationship. It replaces the arbitrary managers with scientific and rational processes but the limitation of that theory is to ignore the human nature (Weber, 1964). Modern organizational theories focus on human-environment perspective and the emphasis shifts from point of individual, local rules to systemic research and organizational environment (Duncan, 1972). The organizational theories present can establish a good basis on a continued improvement in business.

DT and IM concept is introduced to have a general understanding on technology overview and how it can affect the business performance to some extent (Morabito, 2013). It is a way to develop the organizational performance that includes improving organizational operation and strategic performance. Morabito (*ibid*) also claims that organizations may get great improvements in their operational performance as the investment in DT can result in cost efficiency and improving the product quality. The model of 7R's IM cycle includes the managing information phase where it comes from reading in the beginning to retrieval in the last step (Butcher and Rowley, 1998). The model is a typical paradigm, which can be applied to the digital companies or general companies to promote the effectiveness of managing jumbled information. It has been stated that many companies have to face with information overload due to the information explosion, and most companies attempted to drift towards the information that is highly structured and easy for them to get access with. Thus, DT and IM the theoretical aspects can be seen as a good remind which are attainable for most of the organizations. In this study, IM is rather significant when discussing with Big data to deal with massive information and to reconstruct that information in order to serve for the companies.

Big data in the theoretical chapter is firstly defined with characteristics by using 4V's model (Ishwarappa and Anuradha, 2015). And in brief, Big data has the certain characteristic concerning volume (the amount of data), velocity (speed of data transmission), variety (range of data types and sources), and value (costly efficient). Furthermore, it is Big data which is

identified with its effect in different organizations and business model. According to Stephanie and Barbara (2015), the use of Big data in organizations may address the unanswered problems in order to refine and upgrade their business operation and propel the decision-making. In one factor of improving the organizational model, Big data enables improvements from initiating a functional and strong strategic metric from the commodity based-practice, and it supports the strategic measurement in novel ways by offering new data, insight, and action (Wixom and Ross, 2012). In another factor of the innovating organizational model, Big data can provide rich, exciting chances to leverage and extend an organization's strategy toolbox. Besides, generating corporate value from Big data becomes more reliant upon the strategy makers and surrounding structures and processes in today's environment (Stephanie and Barbara, 2015).

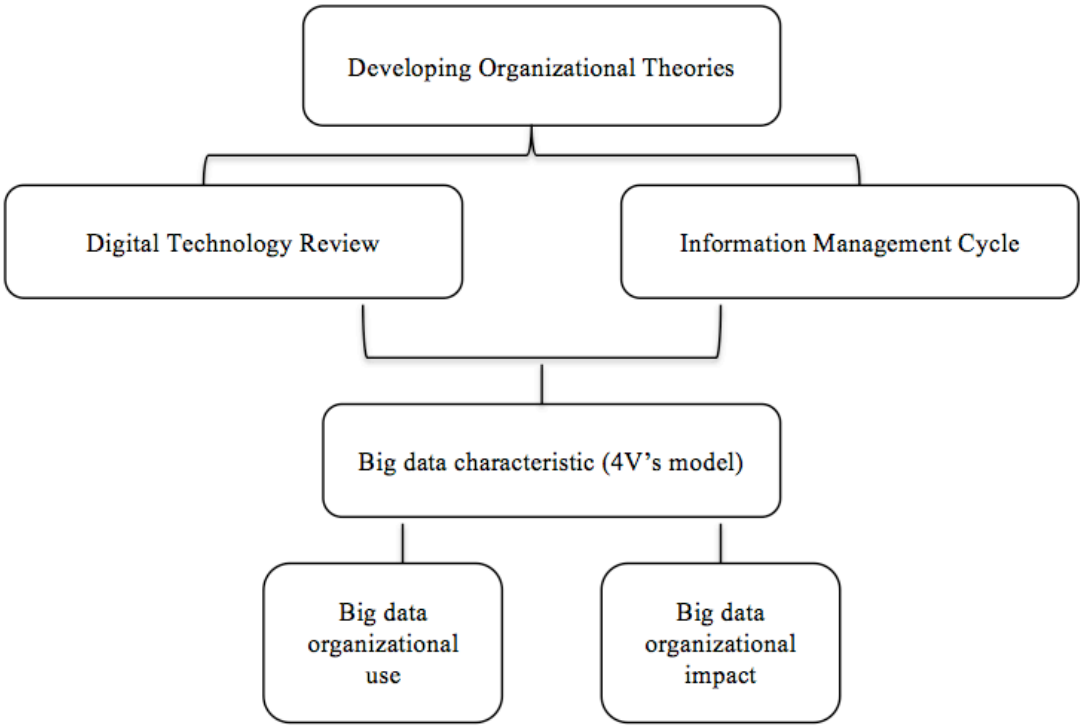


Figure 4. Theoretical framework

Figure 4 depicts how the theories, models, and concepts are used together in this study to obtain a greater understanding of Big data. The organizational theories present the organizational environments and practice, which give the basic explanation factors that are important when improving the way of organizational performance. Organizational theories are a rather broad theory, but quite effective since it can be applied to any organizational field. The framework is bit narrowed since the digital technology review and information management model has been launched. The purpose is to provide background knowledge for readers to have a better understanding of technology effect to business as well as the importance of managing with tons of information. All these mentioned theories are naturally considered to as the cushion effect on Big data. The study on Big data starts with offering the characteristic of Big data which helps know the attribute of Big data. The broad concept of Big data is spread into two sub-features relate to the organizational use and organizational impacts focusing on the organizational perspective. Overall, the concept and the theory together build up with the theoretical frame, which is the key factors in the establishment of

Big data, and present the importance of close interaction and real perspectives in organizational performance.

3 Method

This chapter provides an impression for the study of relevant methodological approaches. And it also measures the study transparency and consistency in order to achieve the research problem and outcomes during the research steps. The study conducts to a qualitative, flexible design, and it is supported by literature review as well as the defined case study with the multiple sources of evidence – semi-structured telephone interview which is considered as the primary sources – also relies on online observations and searches which are maintained as secondary data sources. The trustworthiness and ethical consideration are presented at the end of this chapter.

3.1 Choice of method

The aim of this study is to identify with the current insight on which level the dairy companies in China are initiated with Big data in order to investigate on how Big data can affect the organizational performance. There is an identified gap in a recent study concerning how Big data relates to business activities and how it is implemented in Chinese dairy industry. As a result, this study attempts to bridge this knowledge gap with the initiative purpose.

To successfully achieve the research objective, a qualitative, exploratory study is concerned with interpreting phenomenon in this case. Besides, Yin (2013) gives three different parameters when conducted a qualitative study. It starts with the questions 'what', 'how', and 'why'; and then those questions should establish a good connection with the real-life context. Finally, having less control on the behavior event is also necessary for researchers. In this study, the research question begins with asking to what extent are Chinese dairy companies use Big data. Next it continuous to the usage of Big data and the expected value from which Big data can create. Lastly, all these parameters should be undertaken in order to fulfill and strengthen the choice of method.

3.2 Research approach

A qualified researcher has to have deep knowledge of the specific academic fields that been examined and explored; and such qualification of knowledge have significant impacts on the research theme, theoretical background, and choice of method. As a result, the extent to which you are clear about the knowledge at the starting of the research shows a fundamental problem concerning the type(s) of approach to adapt in the research project (Robson, 2002). Furthermore, both deductive and inductive approaches were portrayed based on the adopted reasoning. In 2010, Ketokivi and Mantere show that the deductive approach occurs when the conclusion is derived logically from the a set of premises; and this causality indicates the conclusion being true when premises are verified to be true. In contrast, when there is a gap which logically exists between the conclusion and the premises, the conclusion being "judged" to be supported by the observation made (*ibid*). Based on these observations, the Big data research topic in this study led us to use an inductive approach in order to generate theories since the primary purpose of inductive approach is to make the empirical results to rise from frequent, dominant, or inhernet of raw data with little constrains from the existing structured theories. The use of an inductive approach is typical in qualitative analysis to build up clear connections between research objectives and findings which derived from raw data and to ensure that relations are transparent and defensible (*ibid*). Furthermore, the use of

inductive approach focuses on developing potential theories about underlying structure that are transparent in text data. In this study, there is little existing literature covering the area of using Big data in dairy industries especially in China. Hence, the purpose of using inductive method is to generate thories instead of testing the existing theoreis by using a deductive approach.

What can be observed in this study is how can Big data, as a tool of DT, affect dairy companies' effectiveness and performance in the upcoming era of Big data. Moreover, a qualitative, exploratory case study is manipulated when assessing the value to which Big data can bring to dairy companies. Data collected in the empirical part is derived from semi-structured interviews, annual reports on defined dairy organizations, website information in different dairy industries, and an extensive literature review on empirical background. Additionally, the semi-structured interviews is considered as a significant section in analysis; the collected information which comes from interviews would also bring information value on the acceptance level for the interviewed companies that invest in Big data. The empirical findings are explored and analyzed based on the theoretical framework, which helps to evaluate the consequences related to set problems, hereby intending to have additional knowledge to the current research scenarios.

3.3 Research design

According to Robson (2011), research design concerns the various things which should be thought about and kept in mind when conducting a research project, figure 5 shows the more visualized framework for the general research design.

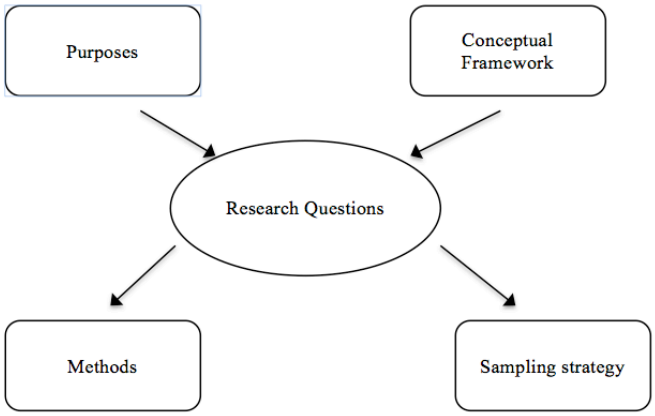


Figure 5. Framework for a research design.(Robson, 2011).

This study attempts to solve a complicated and rather underdeveloped scenario of technological- based business concerning dairy industry. Big data is specifically considered to provide the opportunity for organizational change and information transformation in the information era (Halevy *et al.*, 2009). On the other hand, this study consciously aims to present a real-life context instead of doing speculated experiments to deal with the research problems. According to Robson (2011), "*A flexible design have traditionally included the researcher and the relationship within the boundary of what is examined*". Briefly said, flexible design is much more 'do-it-yourself' and it has the advantage to researchers to reconsider constantly the research paradigms as well as their mind framework, which thereby can result in fresh findings.

The study with a flexible design, on the other hand, needs to be framed within assumptions and some important characteristics such as an evolving design, the presentation of multiple realities, the researcher as an instrument of data collection, and participants' overviews, etc. (*ibid*). Doing architectural framework is essential to the research quality since the flexible design is heavily hinged on the perception of researchers as well as their brain emotions; therefore, the interpretation of empirical data is various, and the threat to giving the right explanation becomes evident (*ibid*). Besides, a study from Uma and Roger (2009) has elaborated that an exploratory study was undertaken when there were not sufficient knowledge about the situation at hand, or little information was available on how similar problems have been solved in the past. In essence, an exploratory study commenced to better understand the problem nature because very few studies have been conducted in that area. In this type of study, the theoretical term of Big data is wide-spread in online sources, and digital technology is becoming one of the crucial value to the business R&D. However, the use of Big data in food industry, particularly in the dairy sector, still lacks researchs and constructive findings from previous works of literatures. Thus, an exploratory study is set up in this study and data is collected through the observations or interviews.

3.4 Literature review

Having a rational awareness on the literature review from this study is a basic step to obtain confidence in studied terms and issues (Yin, 2013). A literature review can be described as to identify, evaluate and critically assess the chosen topic (Jonathan, 2010). Reviewing literatures from the topic dimension can provide an understanding of previous research that is relevant to your study. Furthermore, the essential purpose for building a literature review is to shape and develop the current state of finding on the chosen research area (*ibid*). Having said that, a review on literatures is not limit to develop fixed observation in the chosen areas, as the main objective is to "*concisely summarize the findings or claims"*, that have emerged from previous research effort on a subject (Knopf, 2006). In 2008, Backman listed a three-major-phases with conducting a proper literature review: consulting, internet search, and manual search.

First phase: The topic has a primary selection according to my initiative to explore the meaning, characteristics, and effective use of Big data. Moreover, there is also a theoretical gap between previous research and current scenario of using Big data in Chinese dairy firms. Searching and finding such relevant literatures were suggested by my supervisor. He proposed some advice that the initial review should provide the comprehension on a solid basis from the picked articles or relevant research fields. For instance, the keywords "Big data", "digital technology (DT)", "Big data characteristics", and "Big data management", "Big data company performance" were used in order to locate the academic books or journals firstly. However, a challenge suddenly appears from using database to search peer-reviewed articles and cases of Big data. Previous studies on Big data in business related fields are scarce because the terminology of Big data was proposed in the past ten years and recent articles are mostly focus on informatics, science & technology, engineering, etc. Even so, I was trying to adopt some possible synonyms or relevant research fields to produce a certain range of potential literatures, which attempts to consolidate the theoretical foundation for this study.

Second phase: The next stage comprises Internet-based searching in databases. The main tool was used to access literatures from Primo, which offered by Swedish University of Agricultural Sciences. The databases were also adopted in searching Big data literatures such

as ProQuest, Web of Sciences, Scopus, Epsilon, Science Direct, SAGE journals, Google Scholar, *etc.*, in order to demonstrate a wider aspect as well as a crucial relevance to this study. Nevertheless, the literature search has to be circumscribed from the year 2000 until now when to search for the term Big data (most of the accessible and reviewed literatures are published in recent five years) or other terms such as information management (IM) are not restricted to the creation date.

Third phase: The manual search includes searching books in the library and investigate literature articles from previous studies and relevant works, which were considered as the 'snowball technique' from the reviewing steps (Ridley, 2008). Overall, many books are involved in a well-cited and create specific feature in Big data fields thus the realiablity and trustworthniness have been vastly strengthened. There are published books which directly or indirectly touch upon the use the use of Big data, the Big data value, and also Big data analytics which shapes the solid theoretical framework in this study.

3.5 Data collection

Data gathered in this study is based on qualitative nature from two main sources. The first source comes from telephone interviews with representatives who are chosen from the targeted companies; and the second one derives from the website information and company documents, which are directed to the target dairy companies.

3.5.1 Selection of dairy firm cases

In case study, as elaborated by Yin (2013), the case is the situation, individual, group, or whatever we interested in. The case study should identify each of the selected case itself to a synthetic, connective study instead of doing a sample in one study (*ibid*). According to Eisenhardt (1989), the need for choosing each case is to predict the similar or opposite results contributing to predictable reasons (reproduction of theory). On the other hand, case studies are differentiated with single and multi-cases with various levels, which essentially connected to the data collection techniques like doing observations, interviews, surveys, *etc*. (*ibid*). The potential motives for conducting a case study have numerous benefits because this technique is more effective on building theory, examining theory, or providing descriptive context (*ibid*).

As this study tries to explore the scenario of using Big data in Chinese dairy firms, and as it is known that there are approximately 1,500 registered dairy firms in China (Yujie, 2012). The certain industry comprises a large number of players, and the study requires to have a broad arry of players having willingness to enage the case participation. Further on, the initial targeted informants are proposed since it needs to ensure to what extent of cooperative awareness do these people have which are related to the case study. After that, when the people show a great interest and insight to the first-step interview questions, we continue to introduce a second-step interview which will embrace more specific interview questions concerning their way of doing business by using Big data for example.

After the case is chosen, we primarily focus on the dairy industry, where the main business of these company is to produce liquid milk. The recent information from Takongpao (www,Takungpao.com, 2015) stated that Guoyun Database has firstly displayed the Big data platform for Chinese dairy industry and established a strategic cooperation memorandum with the major dairy firms such as *Yili, Mengniu, and Shengmu Organic Ltds*. These dairy

companies mentioned above have superiority positions on financial capital, working talents, IT support, as well as data availability. Nevetheless, in the current scenario, a data-driven incentive becomes difficult in most dairy firms and it is hard for them to construct Big data and supporting facilities; but competitors from globle environment, grown up in the DT innovation, have achieved considerable success in being more data-driven such as Arla food and Chitale Dairy. Thus, the dairy industry in China is becoming full awareness-driven to be set as a focal point for the research.

The three dairy companies which are used in our case study can be found in Table 1 below. The following criteria concerned with these companies have four listed characteristics: 1) the companies should have the brand icon in their development history – meaning that the companies are selected with their accumulated power instead of instant establishment or investment. 2) the companies have the digital preference indication, which can be found in their production and business activities. 3) Mengniu Dairy has the indications of applying Big data in its operational performance from their annual report and company's website. 4) the companies are represented to have different size and orientation for how their operational activities are exercised. According to these factors, it can make sure that the study has an extensive coverage of the dairy industry in China as well as generalize the similar consequences in international level.

Table 1. Information on selected case dairy companies

Company	Type of company	Number of Employees	Turnover	Year of History	Location
Yili Dairy Group	State-owned	90 thousand	54.43 billion RMB	1993	Inner Mongolia
Mengniu Dairy Limited	Private	50 thousand	43.36 billion RMB	1999	Inner Mongolia
Yiyi Dairy Company	State-owned	1.2 thousand	0.46 billion RMB	1996	Anhui Province

3.5.2 Selection on Interviewees

Yin (2013) highlighted that "*interviews are essential source of information existed in case studies*". The interview has the potential to specify detail information, enrich concepts and explanations of the problem. Especially, in this study, a semi-structured interview was implemented for the purpose of data collection convenience in light of evolving theories and adapting to the respondent of interviewees (Robson, 2011). While it showed that there are quite many positive aspects for a case study, an interviewer should be well motivated and prepared for the interview questions and established high interests and interaction with the interviewee, and also document carefully all possible results (Leech, 2002). In this study, we used the semi-structured telephone interview in order to reach these targeted interviewees, collecting relevent observations and getting responses subsequently.

The four individuals are picked according to the primary selection stage. For the specific reason that representitives need to entirely participate in the dairy industry, from production to selling process, and the representitive should suit to the right position in the dairy company. Therefore, the four representitive should have the right position to be either the company's manager or the people who in charge of the corporate planning and having the knowledge reserve of Big data. In the first attempt, we need to unfortunately redirect one of our representitive in Mengniu, because the first person we contacted was unfamiliar with the company's current status and its willingness of using Big data. In data collection process, it has to be clear that the term of Big data was grown up since the last five years and it has the premise that the new technology field is comparatively new to the dairy industry, as we have mentioned this on the knowledge gap, thus fewer employees from the company would have sufficient knowledge reserve on this issue. Therefore, this is the crucial reason for why we must target the interviewees with the right people in right position, and the benefits for chosing right person can provide as much as information to obtain a good understanding to the research questions. Even though, it is necessary for us to fully respect the interviewees' willingness that if they agreed to put their name on the table or they expected to be anonymous. It needs to emphasize that we have selected two different interviewees in Yiyi Dairy Company, and the reason concerns with two objective facts. On the one hand, Yiyi Dairy is identified as the regional dairy company, the company scale, serving range, and supporting facilities in Yiyi Dairy are incomparable with Yili and Mengniu Dairy. Therefore, it needs to have more people taking part in the interview to ensure the reliability of this study. On the other hand, the CEO in Yiyi Dairy has declared that the company has not applied to use Big data at the moment nor he is very familiar with the new technology term. Thus, we require to have another representitive having more relevant knowledge on this study field. The table 2 below shows the selected interview in accordance with the presented dairy companies.

Table 2. The selected interviewees on dairy companies

Name	Position	Interview type	Duration	Interview Date
Mr. Gao Debu	Company Director (Yili)	Phone	45 min	2016-04-13
Mr. Yao Haichao	Senior R&D Manager (Mengniu)	Phone	40 min	2016-04-18
Mr. Wu Minglou	CEO (Yiyi)	Phone	20 min	2016-04-29
Mr. Luo XXX[1]	Representitive in sales department (Yiyi)	Phone	25min	2016-05-08

3.5.3 In-depth Interviews

From the lists of interviews that conducted in the first phase, the questions are created in order to get access to the theme of research questions. The interviews are subsequently undertaken

[1] Mr. Luo wishes to put his first name anonymous

in a semi-structured manner, from where we directed these questions as the initial aim to focus on, but then different interviewees would show and expand their knowledge upon interview proceeded (See Appendix 1). A semi-structured way of conducting interview has the benefits since interview questions could be prepared beforehand and also encourages the bilateral communication from those people being interviewed may have the motivation asking questions to the interviewer. A semi-structured way of the interview is chosen in this study because it would be easier to compare the cases. Additionally, it was also convenient to revise interview questions in the whole process in order to expand the relevance between topic and questions.

Next, The interviews are arranged with telephone way according to the interviewees' time preference. There are two different interviewees who have been selected in one typical dairy firm because Eisendardt (1989) made the suggestion that two observations usually have advantages than only one and it is also possible to explore how different roles interpret with the one question. The duration of the telephone interviews lasted approximately 25 to 40 minutes and they are held in Chinese. These telephone interviews are recorded throughout the entire conversation in order to seize comments and some keynotes. At the end step of each interview, the information is transcribed and stored so that the collected data would be protected for the convenience use of next step analysis.

3.6 Data analysis

Glaser (1992) defines qualitative data analysis as "*any kind of analysis that produces findings or concepts and hypotheses, as in ground theory, that are not arrived at by statistical methods*". It was observed that qualitative data analysis tend to be exploratory in nature. The nature occurred that once we have gathered a huge amount of data, yet it would be bewildered to consider what to include and what to discard from these data. At this point, it is fundamental to have a clear approach, which would support to analyze the data in an effective manner. In this study, the approach of ground theory is applied in order to get access to data analysis. And the term *ground theory* is the systematic developing of theory from data, thus the term itself is synthetically obtained from social research (*ibid*). From Figure 6, Pace (2004) proposed the five-stages of developing a ground theory.

Figure 6. Five stages in ground theory (Pace, 2004).

According to the figure, we start from data observations; the observation stage helps to identify similar characters that we were capable to build the theory base. After that, we come to use the open coding paradigm which means we had filtered, identified, and then clarified similar words or contents in the database. Meanwhile, in order to form the relevant theory, we attempt to check if any special features within each case, and then to check if these phenomena prevailed throughout in the database. Simply said, it is the step of doing case-by-case analysis and then we integrate these small cases and conducted a cross-case analytical approach. The purpose of doing cross-case analysis is to identify the appealing impression and to further detect if these phenomena are displayed as remarkable in the case samplings. In

the process, we also use the selective coding from the real practice, which means that the similar words and concepts are identified and selected from the database. After that, the results are categorized into different concepts according to the proposed theories we used. Additionally, we also provide with the triangulation technique by connecting our findings from interview cases with other relevant sources.

3.7 Quality Assurance

This section presents the quality assurance in this study. It includes ensuring quality on conducting telephone interviews foucsing on time efficiency, arrangement flexibility, and data consistency. Next, the study shortly discusses the validity and reliability on doing interviews and evaluating empirical results.

3.7.1 Ensuring quality on telephone interview

Sturges and Hanrahan (2004) observes, "*To adopt telephone interviews in qualitative research is fairly common, due large to consider of whether telephone interviews are well suited to the investigations*". Holt (2010) suggests that most of the telephone interviews to be left for the group of qualitative researchers or, at best, semi-structured interviews would be suitable to generate narrative data which has been left in an unexplored field. The practical advantage for this study when choosing telephone interview as the main technique to collect empirical data are listed with three main characteristics.

- ***Time Efficiency***: It is relatively unrealistic, in this study scenario, to prepare the face-to-face interviews, which would be commenced on dairy companies in China. And we budgeted the high traveling costs, and perhaps there is a lower-response rate for the interviewees on face-to-face interviews. Conducting telephone interviews, however, have the unquestionable time efficiency when interviewees and investigators are separated by geographical distance, thus making face-to-face interviews would be comparatively time-consuming.

- ***Arrangement Flexibility:*** Offering the flexibility of telephone interview for informants to have an appointment, which not be restricted by the constant schedule, can increase the interviewee's willingness to cooperate and response. Meanwhile, setting a flexible time for the top-managers may also help avoid risks and uncertainties due to the accidental affairs. Top-mangers usually have floating schedule and conducted with a telephone interview would inherently increase the response rate.

- ***Data Consistency:*** The use of telephone way generates qualitative data at a reliable and consistent level thereby resulting in data analysis to stay at the same level as the process we collected on text (Holt, 2010). In this study, the data is trusted on respondents they shared as the context from which to conduct empirical results. Indeed, staying at the data consistency could be considered as a superior approach, particularly in realistic discourse study.

3.7.2 Trustworthiness

In research practice, it is vitally important to achieve validity and reliability throughout the whole research procedure (Robson, 2011). The use of validity and reliability are relatively general in qualitative research paradigm. In a qualitative case study, it is not easy to seek for the universal, broad theory of Big data in Chinese dairy industry and therefore, the main

interest is to obtain a deeper understanding of the performance of each case we picked for the interviews. This study has undertaken to build upon reliability by using information from different channels, in terms of answer from respondents, and by clarifying limitation aspects in the research design process. In data collection, information evidence was used to ensure triangulation and to improve the trustworthiness. As Robson (*ibid*) suggested that the researcher can avoid confusion and information distortion if various sources came up in the same time and concerned with the same context. This also resulted in constructing a high-effective model, which has been used for the empirical findings as well as future analysis.

The initial purpose of maintaining both validity and reliability criteria is to evaluate whether the empirical results are believable and to verify the accuracy of the study nature (Bryman and Bell, 2007). To fulfill such requirement, we combined inductive approach with the flexible design since the problem might occur with the case study and it probably fail to portray the overall study scheme because the number of participants were still relatively low. In order to avoid this scenario, all of our selected informants were encouraged to have a relaxed and freely chats in the pre-conversations and then, we gradually introduced the interview questions and with some informants showed higher motivation we would somehow extend the topics from where they felt as interesting. Besides, we recorded, and transcribed all the informants' information in case there are any faults when we conducted the calling step. All of our informants got the permission to retrospect their interview record. In addition, the way of our coding and cross-case analysis were compatible with the broad scenario of the usage of Big data in Chinese dairy industry as well as the deeper, cross-comparative reasoning on our findings. And lastly, our findings were triangulated by using various sources. All these implemented methods would lead to increase the reliability and validity of our research (Guba and Lincoln, 1994).

3.8 Ethical considerations

Ethical consideration is considered as the important factor through out the entire research process (Guillemin and Gillam, 2004). The people who get involved in the study have a certain degree of pressure and if the study was proved to be less ethical, the participants would be largely harmed (*ibid*). According to Robson (2011), it is the researcher who needs to make full responsible with ethics and the nature of the research. He claims that it is unrealistic to always inform the participants in advance, but if the researcher realizes he or she needs to have the authorization to ensure the ethical aspects, it is necessary to have consent to the participants in the first step (*ibid*). Ethical consideration in this study are split into two parts: the first part associates with the research process and the second part is about the research field. To ensure the former, all interviewees have made the agreement to be taped through the entire interview process and are informed about the purpose of the study in advance. Moreover, the target interviewees have the absolute willingness to chose in order to show their full name or keep as anonymous in this study because we pay fully respect the private life from the interviewees. Their full names and contact details are collected and kept with their consent in case these are requested to ensure transparency and quality.

Big data may involve with a number of ethical considerations in this study field. As it is mentioned before, Big data is a digital technology which can collect oceans of data that consequently affects the human behavior, and thus will touch upon the sensitive topics as privacy and safety, human choice *etc*. (Science, 2011). The most critical affair is a possible misunderstanding of this term. For instance, when it is applied to the dodgy people to steal individual's privacy, or by cut-throat competition to harm the market principles. This creates

a challenge for government, organizations, and individuals since they need to reassure that a feasible, applicable measure would be issued in order to ensure the right of people.

4 Empirical Findings

In this chapter, empirical findings concerning the case companies are presented. They are derived from the conducted informants and relevant study documents. There are three case companies including four respondents have participated. The step begins with introducing the background information of each selected case companies. Next, the study focuses on the explanation of to what extent Big data are implemented in the dairy companies, and how Big data are suggested to use in each case. Meanwhile, the proposed outcomes, which generate from Big data application in these companies are also demonstrated.

4.1 Introduction to the case study

Since the aim of this study is to identify with the current insight on which level the dairy companies in China are initiated with Big data in order to investigate on how Big data can affect the organizational performance. As the study indicated, the empirical result contains three different sources of data with literature, Internet sources, and interviews. From chapter 3, we know that literature and Internet sources are basically secondary data where they come from previous sources. The literature in the empirical study is mainly derived from almanac book for Chinese dairy in 2012 (Gao, 2013), which objectively records the dairy status in China including most of the large-scaled dairy companies in each year. The literature provides the good understanding for knowing the company's facts such as the scale, financial profit, and cost, *etc.* The Internet source which comes from the company's official website, as well as the company's annual report, provides valuable background information in each case. It is fortunate that from reviewing online information with three dairy companies, we have some hints for these companies' interests of using Big data, which in turn serve for the interview results. The interview data, which is the primary source in this empirical finding, is collected from three companies with four different people. The reason is clarified in chapter 3, as the Yiyi Dairy has smaller company scale, and unfortunately, the CEO has quite limited knowledge of Big data. Thus it is necessary to ask for another interviewee who is familiar with such knowledge. The sub-sections below present the company background together with the company's willingness of using Big data.

4.1.1 Yili Dairy Group

Yili Dairy Group is the biggest state-owned dairy company in China, founded in 1993. The main business of the company is to manufacture and process milk and its relevant products, including milk powder, ice cream, sterilized milk and yogurt with its headquarter in Inner Mongolia in China. It has also grown up in the global top 10 dairy companies that are the only one Asian dairy company in the shortlist (www, Yili Group, 2014). As the industry champion, Yili Group relies on providing healthy products, operation, and vision. And its vision is "*Be the most reliable healthy food provider in the world*", which can be interpreted as being reliable towards all of the dairy channels containing customers, brokers, global partners. It ensures Yili's harmonious development; also leads to creating healthy and high trustworthy ethos (*ibid*).

The construct of the data eco-system in Yili can be traced back to 2008 and to be as one of the most famous dairy companies in China, Yili has made it a priority to complete a product tracking system (www, Qinbei.com, 2015). Hence, dairy consumers can use smart phones as well as the social platform to easily get the information concerning production date,

ingredient list, and expiration date of milk. At the same time, Yili also enriched with information which consumers are more caring for, such as purchasing comments, user instructions, and quality report in each can of milk powder. In fact, the tracking system is one of the reflections on the usage of Big data (*ibid*). From Yili's expectation, Big data is not the mysterious technology; in other words, the advent of Big data era requires each of the individual to show greatest interests and each person can be the manufacturer of Big data.

4.1.2. Mengniu Dairy Company Limited

Mengniu Dairy was founded in 1995 with its headquarter in Inner Mongolia. The major business of this company is to manufacture and distribute dairy products in China. It is also one of the leading dairy product manufacturers in China and its diversified dairy products include liquid milk (such as fresh milk, milk beverages, and yogurt), ice cream, and other dairy products (www, Mengniu.com, 2016). Following with Mengniu's internal division of adjustment, the company has set up a series of reformation, which promoted the synergies among different branches. Unlike the development path from Yili Group, Mengniu is consistently trying to search for international cooperation and trade partners with the deepening of consolidation and internationalization trend at present. In 2003, Mengniu has successively built upon the bilateral strategic cooperation with Modern Dairy (Hong Kong), Danone Group (France), and Arla Foods (Denmark). The international cooperation gets the instant benefit that according to the data revealed from Nielson, Mengniu ranked in the first place in terms of the entire market share with fresh and liquid milk in China in the second half of 2013 (www, Chian Mengniu Dairy Company Limited, 2013).

As the pioneer of the dairy industry in China, Mengniu has the clear vision of the digital technology impact. And the company has subscribed the "food safety cloud" platform agreement in 2015. The intention is to exert the merits of Mengniu's brand, management experience, and information in order to frame the Big data platform on food safety (www, Mengniu.com, 2016). The vice president Wang has stated that Mengniu has achieved a milestone that all of our meadows, dairy plants, and market is stepping into the digital information era; and the company can upload about 2 million datasets from its 50 dairy plants into the Cloud every day.

4.1.3. Yiyi Dairy Company

The 30-year-old Yiyi Dairy Company, abbreviated as Yiyi, was set up in 1987 in Anhui province, where the company has nearly 60 years experience in dairy's farming and processing. Even though Yiyi such has small company scale, limited financial turnover, and popularity, Yiyi is still regarded as one of the most renowned dairy company in Anhui province since the company has consistently concentrated on providing fresh and safety liquid milk. At the same time, the free and fast delivery service from Yiyi becomes one of the successful coporate elements, which are markedly favored by the consumers (www, Yiyi Dairy, 2014). By the end of 2014, the company owns approximately 12 thousand Holstein milk cows together with 40 thousand tons annual production volume in fresh milk. The CEO Wu interpreted with Yiyi's mission " *We want to craft our dairy brand delicately with the unique characteristic, and to provide the priority for milk choice for customers*". In addition, Yiyi has received many high-quality titles such as the example milk base, the A level concerning green food, etc., which is deemed as a great honor to this company and also consolidates the company's mission.

In accordance with the business principle "professional focus, integrity efficient" and having high performance in the drastic competition and challenge in China, Yiyi strives itself to become a high-efficient dairy firm with regulated management, diversified products, environmental friendly operation and highly responsible attitude (*ibid*). Nevertheless, the development concerning digital technology, Big data, and cloud service are not obviously displayed in Yiyi since there are fewer traces from secondary sources concerning how the company relies on Big data to develop its business performance in the future. The following interview intends to solve this perplexed issue from the standpoint of the CEO as well as the one representative from its sales department.

4.2 Summary of interview findings

Table 3 presents the summary of interview results, which collected from the three dairy companies. The table simplifies the key elements with the degree of Big data usage, the benefit or hinderances concerning the use of Big data, and finally the how Big data affects the organizational performance typically on decision-making.

Table 3. The summary of the interview findings

	Degree of Big data usage	**Benefits/ Hindrances**	**Decision-making**
Yili Dairy Group	Medium. Showing a strong motivation on developing Big data. However, the development stage so far is quite low.	**B**: Basically showed from the transformation of traditional channel to high effective channel; Getting rid of mass production, *etc*.	-Generate new type of customer solutions - Change the conventional way from intuition and fewer references
Mengniu Dairy Limited	Medium. Starting using like 'machine-produced data'; having relatively completed digital devices such as base stations, GPS tracking and Cloud; shows a great potential on developing Big data	***B***: Impacts on 'machine-produced data'; expanding customer dimensions, exploring potential users. ***H***: User's privacy problems, immature using stage.	- Increase the revenue through development of machine-produced data - Acquiring more management solutions from software companies - Activating solutions for customers
Yiyi Dairy Company	Very low. Extremely low motivation on using Big data, having limited company scale, funds, and customers. However, it not completely resist on Big data application in the future	***H***: More complicated on management, controlling; unpredictable profits, high risky; safety issues; business environment scenario	-The decisions would be processed under the future business environment & competitors - Providing more relevant and individualized solutions

The table demonstrates that the degree concerning usage of Big data in Yili and Mengniu is the level of medium so far. The two companies have show great interests and confidence with developing Big data in order to strengthen their competitiveness. In Mengniu, the company has entered into the starting stage regarding the use of Big data. They have implemented the

so-called "machine-produced data" and a series of activities such as deploying the use of GPS and Could shows the greatest motivation of using Big data in Mengniu. Compared to Yili and Mengniu, Yiyi unfortunately has a very low motivation on using Big data so far. Lots of limitations such as company scale, funds, and geographical serving range have considerably restricted Yiyi to proceed on using Big data. However, as a state-owned dairy company having long historical precipitation, Yiyi never give up for developing Big data in the future, the company is waiting for the suitable time, opportunity, and market change. The benefits can summarize from the interviews having listed features: promoting dairy production channel to more effectiveness, eliminate mass production and reduce production waste, expanding customer dimension, and exploring potential customer demand. In addition, the hinderances which bings from Big data are particularly involved with personal privacy, safety problem. And for business, the using of Big data still has unpredictable profits which may be risky to disturb the normal function of the organization performance. From the empirics, the decision-making in organizations is considered as one of the key factors to organizational performance. Thus, we choose to have decision-making for the results and it reflects that the three dairy companies have generated their new insights on how Big data can affect decision-making in the organization from generating new solutions, change the conventional way of thinking, resolving the management solutions by illustrating facts and figures, *etc.* All these new insights are discussed in the next interview section.

4.3 Finding from interviews

The next paragraph displays the finding from interviews of the three case companies on what extend do case companies apply Big data in their business activities and dairy enterprises to address research questions.

4.3.1 Case: Yili Dairy Group

"*In fact, Big data is fundamentally useful. We knew that people are used to get statistics as well as consequences from sampling, but today Big data can congregate all the samplings we need in order to have some unlikely acquired phenomenon*"
- As reflected by Gao Debu (Pers. Comm 1, 2016), quoted directly from the telephone conversation.

Mr. Gao Debu is one of the company directors who takes charge of Yili's business in Eastern China. According to Mr. Gao (2016), and relating to Yili's manufacturing and processing business. He also elaborated the necessities and importance of utilizing Big data stands on the following characteristics. Firstly, he mentions that "*Big data is like a maiden voyage in our company, many staffs they talked about it, but they rarely know how to use it*". He also stresses that the tracking system does make consumers get easier to track specific information or items, and this technique has helped Yili collect billions of data from consumers, the tracking system is the prologue to construct Yili's Big data ecosystem; however, the system itself is still not enough to establish a complete data ecosystem. In his viewpoint, the tracking system is purely a third-party tool, which can only collect information of users. And the technique is fairly easy for other dairy companies to imitate. Next, he confirms that Yili was encouraged to establish an integrated, sophisticated system of Big data from the national support and R&D center in next five years. It is obvious that an increasing number of companies are intensively realizing that DT, Big data would be more important to enhance the company's competitiveness and help create extra profits on the completely new business fields. However, he feels that the essential problem about how to exert the entire functions of

Big data in practice – given complex difficulties ranging from business strategy to technical empirics – seems to be an insurmountable barrier for most of the companies especially if they are small-scaled or lack financial support. In addition, Mr. Gao (2016) also maintains that since "*Yili classic milk series*" has acquired a huge success in dairy market, Yili plan to start organizing one of their Big data platform to the company's most classic milk brand. "*I believe our classic series can make a breakthrough without any difficulties*". Mr. Gao explained that the superiority brand is more likely to get in touch with wider consumer groups hence it can generate Big data from a broader geographical location. On top of that, the so-called "brand effect" can help deploy the composition such as milk taste, packaging style, and weight. Therefore, the data they collect would be qualified with Big data variety and veracity.

Furthermore, Mr. Gao (2016) proposes some benefits that can be brought from Big data in Yili dairy. He has then shortly presented the conventional dairy channel, which compared with the new channel after the industry uses Big data. Figure 7 below represents his comments on which was reported from the telephone interview. "*You cannot imagine how big and complex is when it comes from the traditional dairy channel, and we had each well-organized order from raw milk procurement to the final retailing*" as acknowledged by Mr. Gao (2016). From what Gao wanted to highlight comes from the dairy warehousing stage, and he clarified that dairy products are certainly fast moving consumer goods, which means that they are different from clothes, furniture, and other real estates. Mass production in traditional mode has inevitably contributed to capacity waste, especially in fresh milk products. "*Mass production has once led to about 30% milk waste in Yili, it was a tough mission to coordinate the storage and retailing stage.*" There is no doubt that the conventional dairy manufacturing and process mode led to excessive inventory in terms of mass, non-differentiated production. Mr. Gao (2016) continues to explain that what may occur after Yili applies Big data in the manufacturing and processing stage. He affirmed that the fourth step "order processing" would turn into the most principle part thus Yili can use Big data to conduct the analysis in each order; after that, dairy classification will bind up with the manufacturing. "*You will amazingly find once we would accomplish the binding between manufacturing and different dairy classification which base on orders we used the Big data analysis, then we do not need warehousing and wholesaling sections in the entire channel*". Lastly, Mr. Gao predicts that the use of Big data would bring a revolutionary impact on Yili's production channel. He said that the conventional assembly line production with the non-differentiated dairy product would be fundamentally replaced by high-level of information management and analysis technique. Consequentially, the intellectual and personalized production line been equipped with DT and Big data will essentially alter the conventional pattern of dairy industry.

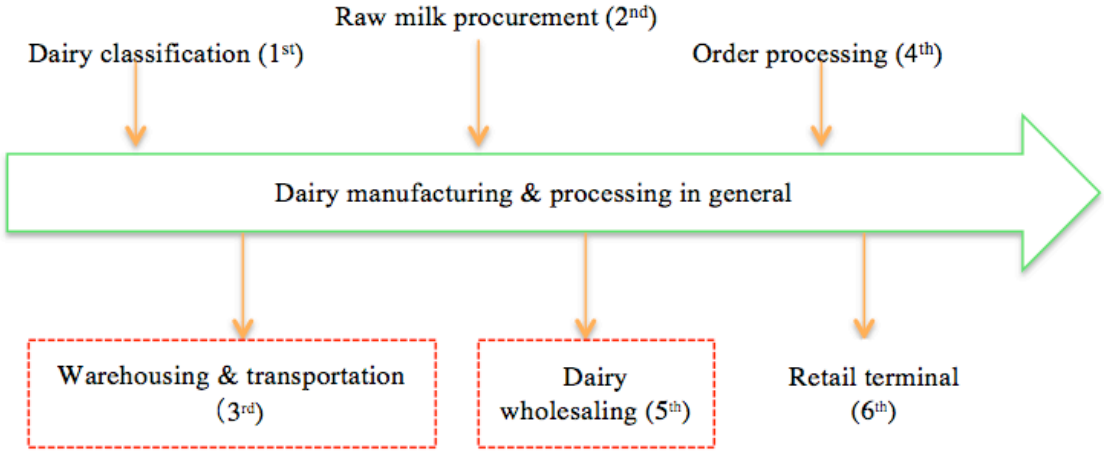


Figure 7. The assumed transmission of dairy manufacturing and processing channel under the use of Big data. (Pers. Comm 1, 2016, as it is created by the author).

The way Mr. Gao sees that Big data can offer numerous opportunities on transforming how Yili conducts business in the primary stage, and thereafter having deeper effects on the company's decision-making. "*If you take this into account, you will feel no surprise that we are attempted to collect all the footprints based on our customer's behind, and then trying to conduct some data analysis in order to promote our sales as well as cultivate the customer loyalty*". As Mr. Gao mentioned.

One obvious effect comes from Big data on Yili's decision-making is the way the data gathered by the s intellectual devices in the company can act proactively towards potential customers. Gao points out that the effect of Big data on decision-making is intensively reflected in two different groups. On the one hand, Big data can provide abundant original and first-hand sources in order to help the get access to massive data from each corner of Yili. In conventional decision-making modes, decision makers are more likely to rely on their intuition and self-judgments due to data scarcity. However, Big data can easily solve the way on acquiring data thus will provide the statistical support to easily handle issues by exporting truth, facts, and figures. With technological upgrade on Big data, massive data through internal and external environments will be collected, analyze, and then exploring the value concerning such data. Therefore, with the help of Big data, Yili has the stronger capacity to predict market demand, conduct precise strategic analysis, and thereby formulate a more effective plan. On the other hand, Big data is able to improve the decision-making quality for Yili in the longer period. Gao believes that the right decision is heavily determined by the data and information quality, and high-quality figures can avoid many uncertainty factors during the decision-making process. Big data technology provides a perfect connection between qualitative and quantitative decision-making.

Lastly, Mr. Gao proposes several opportunities in relation to Big data and refers two main stands in Big data. The first stand appears more close to customer interaction through Big data analytics. "*Customer interaction is something we take it like in-product communication to acquire better responses and personalized messages*" (Gao, 2016). Another stand Gao covers is the development of marketing solutions that offer more directed and dynamic advertising. He agreed that To some extent, Big data can realize the purely accurate market position as well as advertising, because we will completely grasp all of the information (Gao, 2016).

According to Gao Debu, implementing Big data into Yili's dairy business and its relevant process still takes time for company staff to adapt and getting each individual to participate in that game. He clarifies that the companies are trying to make the effort in all different types of directions right now. "*However, I can tell you this is a difficult plan to get Yili on broad. It is not just a round-table conference, and the CEO announces 'ok, let's do this'. It is as much effort to unite our employees to truly believe what we are doing and to follow*". He also believes that technology matters are no long barriers in Yili, but as he said in the beginning conversation, his high expectation on Big data is not mysterious and unrealistic, and each individual can actually use it and make it as part of their daily routine. As Mr. Gao concludes.

4.3.2 Case: Mengniu Dairy Company Limited

"*Big data is, of course, a today's buzzword. We are really satisfied with our digital database (in order to) track and see what people like most in one or more of our dairy products, rate of buying conversation, and how they intend to switch to another dairy but that's all for us to do for the business, we absolutely not attempt to pry into their privacy*"- Yao Haichao, the Senior R&D Manager in Mengniu Dairy Company Limited (Pers. Comm 2, 2016).

In March 2007, Mengniu has initiated with its first R&D center with the investment of 1.2 billion RMB in Inner Mongolia (Liu, 2011). Mr. Yao was appointed as the senior manager in Mengniu R&D center since 2011. According to Mr. Yao, the milk scandal occurred 2008 in China has truly made the company that committed to conducting lots of rectifications. In the R&D branch, the vision of putting the use of Big data analytics was higher up on their agenda in 2011. Yao is proud of using information system and technology in Mengniu's business. He expresses that the purpose of establishing a digital information management system is to react changes immediately and meet the internal control more effectively. Today, Mengniu has applied their information management system on its meadows, dairy manufacturing plants, and transportation system. "*We deploy base stations as well as GPS, which aim to track and monitor every single detail in our procurement, production, and processing channel*". In fact, this is a typical machine or sensor-generated data and it provides real-time data with the massive volumes. Big data is primarily born on the basis of countless generating volumes and also the lightening speed. "*I guess everyone knows 4V's if you get familiar with Big data. And if refers the term 'big', you need to clear that it does not just mean absolute quantity concerning data itself. My understanding is inclined to the diversity and complexity of collecting data*".

Mr. Yao furthermore clarifies that digital tool such as tracking device, base station, and GPS are the examples of the external device to collect so-called 'machine data'. Machine data, as its name suggests, is produced data in an up-to-date scenario with nearby software and electronic device (Yao, 2016). All the relevant applications, servers, network devices, sensors, and other supporting systems, which participate in the entire information management system are continuously creating machine data relating to their status and activities. Thereby, Yao describes the machine-produced data in general thus key information is illustrated and sketched according to the figure 8 below. As it can observe that machine data is produced in a multitude layers with different events and backgrounds, as the software application records and creates data in terms of their specific use. Mengniu has relied on its digital platforms and devices serving millions of users per day will generate gigabytes of machine data that would provide the significant impact into their business operation.

<table>
<tr><th>Event</th><th>Machine-Produced Data</th><th>Business Impacts</th></tr>
<tr><td>In customer's shopping cart</td><td><table><tr><td>Product</td><td>Activity</td><td>User browser information</td></tr><tr><td colspan="3">IP: ***.***.***.***
Cart: "Mengniu QQ Star"
Customer ID: 11235.xt
Location:***.***.Beijing. CHN
Last log-in Time: **-**-2016</td></tr></table></td><td>a. Product performance
b. Activate shopping cart
c. Location tracking & Recording</td></tr>
<tr><td>Complete transaction through card payment</td><td><table><tr><td>Customer name</td><td>City</td><td>Product ID</td></tr><tr><td colspan="3">IP:***.***.***
ID name: XXX
E-mail: 1134er@***.com
Address: ******
Product ID:23453458xxxx.xxx
Time:**.**.2016</td></tr></table></td><td>a. Recording preferred dairy products & customer behavior
b. Update Inventory status
c. Real-time product sales tracking</td></tr>
<tr><td>Write into the database</td><td><pre>% connect to a specific database, using specified username and password
connectionA = database('database', 'user', 'password');

% switch off auto-commit
set(connectionA, 'AutoCommit', 'off');

% insert output data into table
fastinsert(connectionA, 'receiver', colnames, outputdata);

% commit changes to the database
commit(connectionA);</pre></td><td>Source of error</td></tr>
<tr><td colspan="3">Source: (www, MathWorks, 2016)</td></tr>
</table>

Figure 8: The illustration on machine-produced data. (Pers. Comm 2, 2016, as it is created by the author).

According to Mr. Yao, the use of Big data is in the development stage in Mengniu. The organization for data analytics and analysis is heavily hinged on Mengniu's powerful R&D ability and excellent coordination in cross-fields of the department. He suggested that Everybody knows that science and technology are the primary productive forces. Thus, Mengniu has made the commitment to letting more technology companies getting involved with the business operation, managing with the value chain in Mengniu, *etc*. In 2013, Mengniu built the partnership with SAP, a Germany software enterprise, which provides digital & management solutions to its dairy value chain. Also, in 2015, the company announced to introduce IBM to complete the upgrade of our tracking system in order to intensify the milk quality and also create a new tracking mode to their customers.

One of the biggest opportunities of Big data at Mengniu, as mentioned by Mr. Yao, is to widely expand their consumer dimensions. Yao constantly believes that if the company can get more active consumers using dairy products and fresh milk better than the company get those people they know them already and keep staying in one milk brand longer. And if Mengniu is able to know which factors would motivate the interests of customers in a subscription to their service and what can trigger their actual purchase; thus, the company can correspondingly take steps or make adjustments in order to target our potential customers. Moreover, Mr. Yao argues that Big data, when put it in an actual usage, will more or less involve with the privacy issue. He claims that it is unnecessary to worry about user's privacy disclosure. He extends with the bank area that once in the early developing stage, people have

the similar anxiety with how much money the staff would know when they counted the money everyday, and they felt their privacy was invaded. However, when the bank industry has grown up into a matured level, people have less or no worry for that issue because the establishment of supervision rules as well as laws when they come step by step, the privacy issue will be protected in a more sound and regulated environment.

4.3.3 Case: Yiyi Dairy Company

"*We are not just to overemphasize on how Big data affects the consumer buying behavior these days, but we expect to see how Big data can influence the entire (dairy) industry and subvert our traditional concept; we are still on the knowing and learning process anyway*"- Wu Minglou, CEO at Yiyi Dairy Company (Pers. Comm 3, 2016)

Wu Minglou is the CEO at Yiyi, one of the largest dairy companies in Anhui Province, China. Mr. Wu is an expert on the agricultural economy, having worked for several years in rural construction fields and he was also nominated as the NPC representative in China (www,Yiyi Dairy, 2016). As the CEO at Yiyi dairy group, Wu has witnessed the tremendous change at Yiyi Dairy. He firmly believes that milk quality is the cornerstone of Yiyi's life and the foundation of the customer loyalty. When it comes to Big data, his illustrations are recorded "*Honestly, I am not an expert on this area, I have just some learning for this new term, and I remembered I attended one Big data forum last year, but we are not using the 'Big data' at Yiyi. As you know, the dairy scale in Yiyi cannot be compared to those China's top 5 dairy firms; we have limited fund and manufacturing plants, and we do not have our R&D center*". Mr. Wu suggests that the cost is not the determinant factor since the company itself determines the unnecessary awareness to use Big data. Yiyi, as it was reflected before, has quite little factory land and it can only serve their customers in Anhui province especially in Huainan city. Besides, Mr. Wu acknowledges that Yiyi has not merely conduct the investment on technology. He recalls that, when the milk scandal (due to the illegal melamine additives) swept the whole China in 2008, which led to the bankruptcy and regroup in many dairy enterprises. However, Yiyi has hardly received any harmful impacts on the milk scandal. In surprise, the company has invested 2 thousand million RMB at that time in upgrading the milking equipment and also improved its distribution system. Customers in Yiyi were really stable with higher trustworthiness and loyalty, thus the companies were not so worried about the sales and profits. Meanwhile, Wu points out that the lack of specialists with Big data expertise today brings great obstacles for Yiyi. He shrugs to say that it is not easy to find strong expert who can rein Big data at present. In Big data area, China is only a follower at the moment and it is unrealistic for a small dairy firm finding one expert who is willing to contribute in this unexplored dairy area. "*Even so, I prefer to have the 'follow-up' strategy which means I would like to see Big data expand in dairy industry for one day and then it should be the right time to 'follow' with the entire industry. Anyhow, I am not willing to become an adventurer for that*".

The 2nd interviewee is selected in Yiyi's sales department Mr. Luo, Luo wishes to keep his name anonymous when he is asked the questions regarding Big data analytics (Pers. Comm 4, 2016).

Mr. Luo holds the opinion that the opportunity for using Big data in Yiyi is immature and underdeveloped. He explained that in small scale dairy firms, there is a strong resistance to switching the current business and management form to a more sophisticated paradigm with the reliance on Big data. So to say when the dairy firm decides to import Big data like Yiyi,

the company has to extract the large proportion of our incomes and revenues to build Cloud base, and to embed smart devices through all data collection processes. Consequentially, it will become a huge amount of expense to the company; however, it is true that nobody can predict if such measurements would bring about immediate benefits. "*We are currently not considering to use Big data in our short-run plan unless our competitors they all get into that field*". On the other hand, Luo reasserts that competitor scenario and external market environment make Yiyi must stick to its plans and policies more cautiously. Mr. Luo also indicates that in today, a lot of customer purchases conducted based on universal factors such as user experiences, product feelings, and transaction price, *etc*. In a best-case scenario, Big data may affect some factors in that the company will not expect to pay for our dynamics on the non-decisive factor. In the future, Yiyi will take into consideration of market environment and how the reaction from the entire dairy industry. Even though Mr. Luo remains optimistic attitude on the impact of Big data to business, and he believes that digital data would gradually become one of the most valuable assets for the company in the next five to ten years. He hopes that the size and financial funds on dairy company will not become the constraint to developing Big data in the future. In Yiyi dairy, the company will never ignore the intangible value, which can be brought from Big data. In sales sector, Big data can be precisely implemented on grasping consumer behavior and serving for customized, personalized dairy product according to their preferences. Consequently, Big data will enormously promote the decision-making ability on the customer insight and market response.

Mr. Luo points out the main hindrances concerning the use of Big data in the end. First and foremost, he mentioned the user safety is the most crucial problem because the user information may be disclosed in any data transmission and storage processes. He said that in China, there is even no technology enterprise like Apple in U.S, can excellently protect user information relying on IOS and Touch ID with billions of user background. It also can be inferred that a disaster may take place once the customer knows their personal information is exposed no matter what the companies think they are deliberate to disclose this information or not. Furthermore, Mr. Luo lists two visible problems concerning information safety and challenges. On the one hand, he proposed that the conventional data protection would be altered since more terminal devices are connecting to database; thus it has higher possibility for user privacy been threatened. Once Big data in a company been tampered or stolen, losses will inevitably occur. On the other hand, information protection on users still has a large gap between China and Western countries. A number of citizens lack the awareness to protect their personal information and in some situation even if their information is leaked or invaded, they have little possibility to realize. "*Both our enterprises and customers have to get ready for these challenges, no matter how big the company scale it is. The Big data era will step into the upcoming event, but still, no one can anticipate when it arrives*" - as terminated by Mr. Luo.

5 Analysis and Discussion

In this chapter, the empirical findings are analyzed on the basis of the relevant theories and literature presented in Chapter 2. The analysis begins with identifying digital technology and also the usage of Big data in case companies, which is expanded with the information management theory. Besides, the analysis is followed by the results connecting to Big data characteristics as well as organizational impacts to the case companies. The research questions presented in Chapter 1 will also be discussed below:
R1) How, and to what extent, do Chinese dairy firms apply Big data in their business?
R2) What are the main benefits or hindrances concerning the usage of Big data?

5.1 Analysis on Information Management Theory

According to Morabito (2013), the association between data, information, and intelligence are closely connected. The empirical results reveal that the dairy companies in China have the clear recognition of technology superiority because Big data can leverage the effectiveness of their manufacturing channel as well as having the quick reaction on changes. Raw data, information, and the use of intelligence are managed in a well-organized paradigm in the case companies. This information management also generates the general vision on how the dairy companies in China take actions on Big data to their business in their certain context.

As the empirical chapter shows, the three dairy companies namely Yili, Mengniu, and Yiyi have given their attitudes towards how they treat the data-generated information. Dias (2001) suggests that information management can be seen as a transformation procedure of data into information, which was operated by the decision hierarchy in an organization. Information management can also bring about instant benefits from improving productivity and quality, reducing operational cost, and allocating decision-making process. In Yili and Mengniu , it is evident that the original and unstructured information are collected with the help of digital system and advanced devices such as terminal track, base station, and GPS, which passively gained from customers and external environment. Even though the company Yiyi has not intended to develop Big data today, they also realize the privacy information and the hazards for the information abuse. According to Butcher and Rowley (1998), managing with information should be organized in a well order and the seven steps, which come from reading to retrieving, are naturally essential. In the case scenario, Yili and Mengniu have demonstrated their strong ability with resource integration. For instance, the information collecting in Mengniu is established through its manufacturing and processing step. Furthermore, the ‘machine-data’, which is showed at Mengniu, has sketched the scene on how these data are recorded and write-in the database. The company has achieved the management technique on fulfilling its designed step. Likewise in Yili, the company proposes to restructuring its manufacturing and processing channel with the use of Big data. The information management paradigm would probably be switched from a data collecting-driven to a data analytical-driven activity. In other words, the step of reviewing may be noticed as the previous boundary since the company sees what happens from the transmitted data. Nevertheless, in Big data era, the boundary would be extended to retrieval step since the information will substantially rely on information collection or repository, and it conducts with some retrieval techniques such as customized tools, business intelligence, and analytical techniques that compose the portal decision processing system.

5.2 Analysis on Big data Characteristics Model

In chapter 2, the 4 V's model of Big data characteristics is developed to understand the certain features on Big data and how the features lead to different paradigms from the data application. Yili and Mengniu are the two typical cases having a strong motivation of using Big data. The empirical findings show that Yili and Mengniu are top five dairy company in China, the two companies have exceptional advantages where they can easier to get access to talented expertise, supporting facilities, national funds, *etc.,* which lead to a better corporate environment on developing Big data. More importantly said, to investigate Big data characteristics in large-scaled company is considered to have persuasive, extensive, and meaningful effects. On the contrary, the small-scaled dairy companies such as Yiyi has a huge gap with Yili and Mengniu, the company shows quite a low degree of using Big data. The most important reason for not taking Yiyi into consideration comes from the restriction of the company itself. Big data analytics should make sure the breadth and depth of data and Yiyi are not qualified with gathering data from the extensive fields due to its limited serving range and geographical location. By using Yili and Mengniu for this analysis scenario, we discuss with the Big data 4V's model which have been presented in the theoretical chapter.

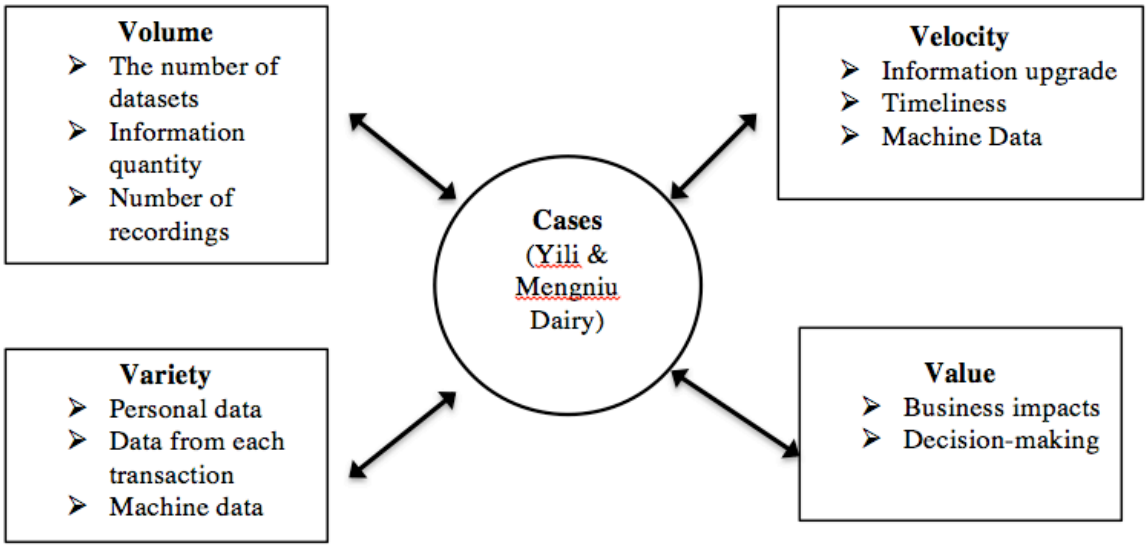


Figure 9. The Big data analysis from case companies' perspectives in characteristics.

Figure 9 is a summary of empirical findings conducted from interviews to identify which factors are included in the Big data characteristics with the detection from Yili and Mengniu Dairy. The analysis has the theoretical basis on the model of Ishwarappa and Anuradha (2015), to show where the theoretical-based features are linked with the phenomenon in reality.

Volume: Hurwitz, *et. al* (2013) expressed the volume of Big data simply means the quantity of generated data. However, the quantity of Big data differs from the conventional sense of the data volume and Big data is usually counted by means of Terabyte, Petabyte with the base point in millions and billions of datum. The case study indicates that Yili and Mengniu, as they are the top 5 dairy leading companies in China, have the full capability to rely on digital devices and the Cloud to collect countless data in each of their business performance. Hence,

it is less challenging for large-scale dairy firms to worry about the issue of volume since those companies grasp tremendously extensive resources in their business activities.

Variety: According to Alexandru (2013), data is not identical and it is necessary to classify into different types, attributes. In Big data scenario, the data varieties are versatile embracing structured, unstructured, and complex structured forms. Regarding the variety of the data sources from the case study, it was stressed that Yili and Mengniu are not focusing on the unstructured data especially the part that comes from the dairy manufacturing process. In Mengniu, the company has come up with machine data when it gets into purchasing and payment procedure. Indeed, the variety of data flows from the shopping cart, receipt line into the database, from which the company will be easier to get information such as customer behavior and purchasing habits.

Velocity: It regards how fast and timely at which new data is created and the speed at which data circulate around (Ishwarappa and Anuradha, 2015). In reality, the data transmission hinges on the consistency and coordination from digital devices. On the other hand, in the information era, to speed up data transmission and its relevant information is no longer a difficulty for modernized dairy companies as digital technology will eventually equipped in the whole industry.

Value: According to Ishwarappa and Anuradha (*ibid*), the value is the most important aspect of the Big data characteristics. The value of Big data is profound and lasting which can be observed in the case study. In Yili, the informant has elaborated the value about enhancing the dairy industry channel. It is clear that the use of Big data can avoid mass production, which can reduce milk inventories and eliminate the wholesaling procedure. In Mengniu, it is also evident that machine data can bring about business value on product performance, location tracking, inventory updating. In all, all the benefits that generate from Big data will become an indispensable support towards their decision-making and strategy planning in the future.

5.3 Analysis on Organizational Impacts

The listed organizational impacts with using Big data, according to Katal *et al.* (2013) and Stephanie and Barbara (2015) propose a good framework to illustrate the organizational impacts in terms of the two following aspects:

- ***Improve the organizational model***
  - New data
  - New insight
  - New action

The publication by Stephanie and Barbara (*ibid*) has investigated that the upsurge in digitalization, accompanied by the increase in data volume, which have prompted companies to reexamine their business activities and also relevant impacts. In the first hand, the usage of Big data can help organizations resolve burning questions in order to optimize their business process and decision-making. In this scenario, Big data plays a supportive role in measuring organizational impacts by offering new data, insight, and action. *New data* occurs when sensors, intelligent devices are inserted to product manufacturing and processing process to track the actual product usage. The original intention of generating new data is to track, explore, and understand customer preference and behavior. In the real cases, Yili and Mengniu have reached the same perspective to upgrade their current devices in order to have more insight on new data. Yili has clearly shown that massive data where the company collects through the internal and external channel will consolidate the company's ability to

predict market demand, conduct precise strategic analysis, and thereby formulate a more efficient plan. In Yiyi, the company still has the awareness new data has the intangible value since it can predict the customer buying behavior and serve customized product even if they are trapped in the problems of company scale and financial funds. *New insight,* as it suggested by Wixom and Ross (2012), has its purpose on identifying outlier to data variation. More importantly, the data transmission needs to depend on visualization software to get the result accurately. The two case companies, Yili and Mengniu, have their ability to establish analytical software since these two companies have put the Big data plan into their constitution in the next coming years. Especially in Mengniu, the document observation from its annual report indicates that the company has cutting-edge technology and by cooperating with the international software companies will make it easier to grasp the advanced technique thus the coming of new insight of Big data is just a matter of time for these two companies. *New action* takes place when companies become well armed with Big data and proficient at making insights based on the generated data (Stephanie and Barbara, 2015). When associated with the dairy industry in China, the case companies represent that Big data is not fully equipped for these industry businesses today; and it was restricted by the time, specialist, and core technology, *etc*. Nevertheless, it is good to see that top dairy companies such as Yili and Mengniu, has taken inspirations on developing Big data in serving for their future business activity and decision-making. In small-scale or private dairy firms, it is suggested to enhance the Big data awareness as well as to improve their horizon field to see what are the performances from international dairy companies and always to keep on learning advanced ideas and technology.

In another aspect, companies use Big data to explore different ways in order to maximize their revenues. Data monetization and digital transformation are two distinct features when we move on the impact on organizational innovation. Digital transformation occurs when companies leverage digitization to go into entirely new industries or to create new ones.

- ***Innovate the organizational model***
  - Data monetization
  - Digital transformation

From the theory revealed, the company scale cannot determine the *data monetization* facing with the case companies. Wixom *et al.*, (2013) ever identified that data monetization is a kind of activity to have the exchange with information-based products to legal tender. Also, regardless company type, data monetization needs organizational group or business unit to build up technical, and business competencies appropriate for the scope of its monetization business. From the results reflected by Yili and Mengniu, they have given a high expectation to the value of Big data regardless of their main business type. It is essential realistic that top dairy companies with data collected from each corner during their manufacturing and processing phase has the monetary value once they sell the data to third-party institutes. Thus, the margins can be achieved from its information business dwarfed the return ratio associated with distribution. Nevertheless, what Yiyi worried about comes from data safety and privacy problem. In data monetization area, data leakage is deserved to ponder during the Big data activity. Personal information and confidential corporate data are vitally important to be conserved and protected in the safe place. It is fundamental to classify data with different features and attribute otherwise data monetization will be not supported in business activity. *Digital transformation*, when examined by Weill and Woerner (2015) that Big data can enable companies to transform from its main business into new industries or ecosystem meanwhile by changing the traditional industry structure. For Chinese dairy industries, Big

data has its main functionality to serve and promote the efficiency relating to dairy's business activities such as procurement, manufacturing, processing, and precise targeting. Furthermore, Chinese dairy industries expect to use Big data in order to guide and providing effective results on helping their business decisions or making business strategies more wisely. As a result, digital transformation is currently less pragmatic in Chinese dairy companies since the start-up phase of Big data in China is still too early and it is hard to transform the dairy business to another new industry field then.

5.4 The Statement of Using Big data Today

As the empirical findings presented in chapter 5, it is evident that the studied dairy companies have their own procedure and plan on developing Big data practices. The first frontrunner, which can be detected among three cases, is Mengniu, as the company has the ripest vision and supporting devices to develop Big data. The company has gotten fully prepared for the upcoming event of Big data since lots of activities such as building the Cloud database, and leading in the so-called 'machine-produced data', which shows a strong initiative on using Big data in their business activities. Yili is also one typical example that shows with a high motivation on adopting Big data into business in its five-year plan. However, with restricted to a series of problems, the implementation of Big data still lays on the agenda at present. In Yiyi, the company shows a low motivation on using Big data in their business due to the small-scale company size, limited service scope, and huge amount of input costs.

Melville *et al.* (2004) argue that technology can support product and service differentiation or the innovation rate in order to promote the company's overall competitiveness. More specifically, organizations will truly benefit from the use of new technology since it may lead to cost efficiency and product quality, which in turn raise the enterprise's comprehensive strength. The study suggests the scenario in Chinese dairy firms, such as Yili and Mengniu are perusing their right direction on adapting their business activities to the new digital technologies like Big data, Cloud Storage, GPS tracking, *etc*. Whereas in other small-scale companies like Yiyi, it has limited ability to deal with Big data and its supporting facilities thus managing with Big data is difficult for these companies in the certain situation.

Morabito (2013) acknowledges that DT can create right solutions to the users together with providing technological guidance as converting them through the industry channel into economic outputs. This perception can be applied when plotting out the scenario in the dairy industry and knowing to what extent Big data can be used in business operations when taking into account of various influential factors. It is conspicuous when to think small dairy firms have not come as far as the frontrunner in the use of Big data analytics because it is less urgent and not enough aspiration to transform their business activities by the use of Big data. Actors such as Yiyi, has an insufficient initiative to use Big data into its real practices. As Mr. Wu Minglong explains in the Yiyi case, he points out the successful roads to a company are different; and Yiyi has its own way to find the keyway to the company success instead of using Big data. On the other hand, his attitudes show that Yiyi does not absolutely resist using Big data in the future since the market scenarios are constantly changing. Other actors, such as Mengniu, does not reveal the same reaction – something that is required to commercialize Big data and capture the value relying on such technology input. Therefore, it can be inferred that the use of Big data is fundamentally hinged on the dairy company's overall strength. Digital technology investment such as Big data can partly help these companies to be more effective and competitive in the current background. Nevertheless, Big data is rather a

determinant to dairy firms that having limited ability and plan to serve for the certain range including customer volume and geographical location.

5.5 Identifying Opportunites of Big data

In this section, the study identifies possible opportunities that Big data enable to the case companies relate to literature.

According to Halevy *et al.* (2009), Big data defines opportunities by which companies can provide real insights on the decision-making and value-to-customers. In this study we have found several benefits and opportunities which brought about by Big data with the connection to how dairy firms can make such measurements: create business values, deliver user values, and provide values for decision-making.

In the first place, Big data can help the dairy companies create business value by improving its technological merits. This is an opportunity that was reflected in all of the case companies, underlining the benefits or potential value from which Big data will bring about. Mr. Gao (Yili) and Mr. Yao (Mengniu) have emphasized being motivated their personal suggestions on using Big data as one of the crucial ways to expand customer dimensions in relation to providing resolution for cutting inventories and alter mass production scenario. Mr. Wu (Yiyi) also held an opinion that Big data will make great impacts on the whole Chinese dairy industry; even though the company has limited ability to develop Big data at present, they intend to see the value in the future from which Big data can generate. These underline a close interaction between the creating values for business and seeing values from the different corner of the dairy industry. Secondly, our findings show that Big data can deliver user value in terms of using machine-produced data to precisely deal with the information from each user. Mr. Yao (Mengniu) explained how the machine-produced data worked in the medium of digital tools and to identify the events coming from the customer's shopping cart to write-in stage. By observing those personalized details, the companies will be more proactive to know factors that can motivate their interests in a subscription to the service and what can trigger the customer's actual purchase. Lastly, it is more explicit that the opportunity on Big data's decision-making as it was shown on Yili, Big data could provide a mass of first-hand information around each corner of the company. The collected, centralized data becomes more convenient for decision-makers to take look up the overall picture at the company. Furthermore, Big data not only generates information based on the data volume; as it can be reflected from Big data characteristics, Big data also offers various types of information which indeed enhance the data diversity, thus, can help decision-makers conduct strategies and plans from different angles. On the other hand, Big data can promote the quality of decision-making with the longer period. It is believed that the right decision is heavily determined by the data and information quality, and high-quality information can avoid many uncertainty factors during the decision-making process. Big data technology provides a perfect connection between qualitative and quantitative decision-making. An explanation of the slow development in accordance with Big data implementation can be found in the Yiyi case. The company supposed that a reason for not to invest Big data can be interpreted as cost issues, lack specialists, *etc.* However, the company remained positive to the use of Big data in the future. It can be inferred that the decision-making in Yiyi will depend on the visual dynamics as well as instant benefits from Big data, which can be predicted in the near future. Thus, we cannot ignore the value of Big data that there are plenty of opportunities concerning a more data driven and Big data implementation into the dairy industry by the upcoming period.

5.6 Identifying Hinderances of Big data

As it can be found above, there are many opportunities created by Big data since it creates business values, delivers user values, and provides values for decision-making. However, it is also evident that small dairy companies in China have currently not come as far in restructuring their plans to develop Big data on purpose compare to the large dairy companies. In this section, the identified inhabitors will be discussed on why this may occur.

When closely link the inhabitors to the empirics, it is evident that Mengniu and Yiyi have come up with the privacy concerns in coincide. Mengniu has brought up this problem once the use of tracking devices and when it transforms data into the data warehouse, personal information will automatically record and appear through the data transmission. Although Mengniu has expressed that the company would absolutely not attempt to pry into personal privacy, the oral commitment seems to be pale towards their customers. What Mengniu could do, as elaborated by the senior manager, is to perfection the supervision rules and since Big data is still underdeveloped, immatured across the Chinese dairy companies. It requires more institutions to get involve in order to regulate the operational principles. Yiyi shows another perspective concerning protect user privacy. The company claims that it is necessary to strengthen the user's awareness to protect their privacy. However, this perspective is considered to be quite vague. It needs to argue that people have their intuition to protect their privacy information; but in Big data era, the situation becomes different and to protect the privacy is not solely hinges on promoting people's awareness. Digital technology offers conveniences the way people live and when we enjoy the technological benefits, it is inevitable to have personal data login. Therefore, Mengniu demonstrates a feasible suggestion to promote Big data into a real applicable field where it has to be supervised and controlled by common standards.

Another inhabitor that brings up in Yiyi is the cost problem. It can be associated with small-scaled dairy firms that most of them have to confront with a high initial cost as well as human capital-intensive operations realized as main hurdles for implementing Big data in business reality. For large-scaled dairy firms, using Big data may regard to be an experiment since the top firms have the particular superiority to allocate resources and trial Big data, which comes with a conspicuous opportunity cost. In spite of this, using Big data for large dairy firms becomes one of the significant paces to be more competitiveness in the international dairy market. However, the small dairy firms do not have the profitability requirement in the same way as the commercial large dairy firms they play. In an overall opinion, this may interpret why Big data is not prioritized in Yiyi at present.

6 Conclusions

The aim of this study is to identify with the current insight on which level the dairy companies in China are initiated with Big data in order to investigate on how Big data can affect the organizational performance. Meanwhile, the empirics also investigate on the pros and cons when using Big data in dairy organizations and how Big data can affect the decision-making which is recognized as the typical example of organizational performance. This last chapter reconnects to the aim and summarizes the key findings.

The research shows that firms are currently at the different developing stage in Chinese dairy industry concerning the usage of Big data into their business. Yili and Mengniu are typically Chinese top five dairy firms, which can be deemed as the industrial pioneer to Big data's industrial application, whereas Yiyi has initiated with low practice but, owing to various problems, not come as far in this developing scenario. We also observed that the increasing levels regarding Big data utilization could be described as the reason with the firm's overall strength and urgency degree to expand business and increase potential values in terms of the Big data revolution, which takes place today.

The frontrunners Yili and Mengniu, they are showing a high level of awareness on developing Big data in their future activities. Both companies have proposed a rough plan on either broad overview about industrial channel or microscope of how to generate 'machine-produced data' inherently proved that Big data could help navigate their performance and as a consequence focusing on the area for value creation. On the other hand, even though Yiyi displayed with low motivation on Big data, the company consistently maintains that they will follow the market scenario once Big data will be applied commonly to the dairy industry. Thus we can conclude that Big data adaptation may visibly or potentially function to leverage the business performance, and accordingly bring added value towards its future growth, although the answers are biased with whether or not the research attempts to solve for Chinese dairy industries, other industry generalization would have to be furthermore investigated.

The study suggested several opportunities, which propelled by Big data with a connection to business activities where dairy firms intend to grasp and deliver value to their customers. These opportunities are particularly reflected in the aspects of creating business values, delivering customer values, and provide values for decision-making. At the same time, the study also identified some hindrances to Big data utilization in dairy organizations. The most obvious obstacle as revealed in Mengniu and Yiyi is concerned with privacy disturbance. And it is inevitable to link with the privacy issue when data generated with countless volume. We have to lay our hopes on solutions such as multiple encrypted databases and more consolidated regulations Nevertheless, it has to be seriously considered that if an enterprise fails to have a foresight in the long-term and only pursuing its current benefit, it will have difficulties to use new digital technologies like Big data to fundamentally alter their business performance. Other hindrance consisted of promoting the Chinese citizen's awareness on privacy protection and preventing illegal data invasion. Finally, the opportunity cost of spending excessive resources on the usage of Big data is stressed as a potential inhibitor for dairy companies with limited resources. The small-scale dairy firms are reluctant to develop Big data immediately may contribute to the fact of lacking sufficient resources. What they seek to find are instant benefits and optimal inputs to ensure the market stability.

The future research can be expanded to other areas like marketing and customer needs, and the research method can apply quantitative study in the next step of Big data research. For

instance, how does Big data add value to know customer needs and improve customized service? Future studies can make more efforts to research other fields.

The second interesting studying field is to discover how to apply Big data analytics approach successfully as the corporate competitive advantage and to develop the key successful technologies. It is difficult to collaborate the current Big data-driven dairy companies, and the future research is likely to know what barriers during the Big data implementation procedure. Since data-driven organizations are increasing and getting more connected with different sensors and devices, it also creates a higher risk for information system to be penetrated by hackers or damaged by break-ins. Therefore, the third interesting future development topic can focus on the high-security risk concerns and the approach to protecting the Big data user privacy.

The last possible research topic may be the change of a company's culture where resistance might occur towards the change. Altering a company's culture will create challenges when implementing a digitized strategy among dairy companies. Hence, consensus throughout the organization is crucial. As the Big data and digitalized strategy growth over time, challenges that dairy firms encounter when they engage in extensive technology usage lack of knowledge to combine all data into a common platform, which may cause barriers in regards to interoperable system and difficulties in integrating different sources of data. Future research can view it as a good point to study as well.

Bibliography

Literature and publications

Alexandru, A., 2013. Big data Challenges. *Database Systems Journal*, 5, pp.31-40.

Backman, J., 2008. *Rapporter och uppsatser*. Lund: Studentlitteratur.

Banica, L. & Hagiu, A., 2015. Big data in Business Environment. *Economic Sciences*, 14(1), pp.79-86.

Barney JB, 1991. Firm Resource and Sustained Competitive Advantage. *Journal of Management*, 17(1), pp.99-120.

Barua, A. Kriebel, C H. & Mukhopadhyay, T., 1995. IT and business value: an analytic and empirical investigation. *Information Systems Research*, 6(1), pp.3-50.

Brown, B. & Micheal, C., 2011. Are you ready for the era of Big data. *McKinsey Quarterly*, 4(1), pp.24-35.

Bryman, A. & Bell, E., 2007. *Business research methods*. Oxford: Oxford University Press.

Brynjolfsson, E & Hitt, L., 1996. Firm-level evidence on the return to information systems spending. *Management Science,* 42(4), pp.541-88.

Butcher, D. & Rowley, J, 1998. The 7 R's of information management. *Management Information*, 5(3), pp.34-36.

Cronin, B & Davenport, E, 1991. *Elements of information management.* New Jersey : Scarecrow Press.

Daft, R, L. 2007. *Organizational theory and design*. Thompson South-Western, Boston.

Dias, C., 2001. Corporate portals: a literature review of a new concept in Information Management. *International Journal of Information Management*, 21(1), pp.269-87.

Duncan, R.B., 1972. Characteristics of organizational environments and perceived environmental uncertainty. *Administrative Science*, 17. pp.313-327.

Eisenhardt, K., 1989. Building theories from case study research. *Academy of Management Review*, 14(4), pp.532-50.

Fox, P.F., 1995. Lactose, Water, Salts and Vitamins in Milk. *Advanced Dairy Chemistry*, 3(2).

Gao, H., 2013. The almanac in Chinese dairy industry 2012. *Beijing: Chinese Agricultrual Press*. pp.11-488.

George, G. & Haas, M., 2014. Big data and Management. *Academy of Management Journal*, 57(2), pp.321-26.

Glass, R. & Callahan S., 2015. *The Big data-Driven Business*. New Jersey: John Wiley.

Guba, E.G. & Lincoln, Y.S., 1994. Competing paradigms in qualitative research. *In Handbook of qualitative research*. Oaks: CA: Sage. pp.105-17.

Guillemin, M., & Gillam, L., 2004. Ethics, reflexivity, and "ethically important moments" in research, *Qualitative inquiry*, 10 (2), pp. 261-280.

Guo, X.G., 2013. The History of Management Theory. *China Economics Publishing House*. Beijing, China.

Halevy, A, Norvig, P, & Pereira, F, 2009. The unreasonable effectiveness of Data. *IEEE Intelligent Systems*, 2(1), pp.8-12.

Holt, A., 2010. Using the telephone for narrative interviewing: A research note. *Qualitative Research*, 10(1), pp.113-21.

Hu, D.H., Reardon, T., Rozelle, S., Timmer, P. & Wang, H.L. (2004). The emergence of supermarkets with Chinese characteristics: challenges and opportunities for China's agricultural development. *Development Policy Review,* 22, (5), pp. 557-86.

Hurwitz, J., Nugent, A., Halper, F. & Kaufman, M., 2013. *Big data for Dummies*. Hoboken, New Jersey, Cananda: Joha Wiley.

Ishwarappa, P. & Anuradha J., 2015. A Breif Introduction on Big data 5Vs characteristics. *Procedia Computer Science*, 25(2), pp.319-24.

Jonathan, W., 2010. *Essentials of Business Research*. London: SAGE Publication.

Katal, A., Wazid, M. & Goudar, R., 2013. Big data: issues, challenges, tools and good practices. *International Conference on Contemporary Computing*, 1(3), pp.404-09.

Ketokivi, M. & Mantere, S., 2010. Two strategies for inductive reasoning in organizational research. *Academy of Management Review*, 35(2), pp.315-33

Knopf, J.W., 2006. Doing a Literature Review. *Political Science & Politics*, 5(3), pp.127-32.

Kros, J. F., Richey, R. G., Chen, H., & Nadler, S., 2011. Technology Emergence between Mandate and Acceptance: An Exploratory Examination of RFID. *International Journal of Physical Distribution and Logistics Management*. 41(7), pp. 697-716.

Leech, B.L., 2002. Asking Questions: techniques for semi-structured interviews. *Journal Political Science and Politics*, 35(1), pp.665-68.

Liang, H.B., 2014. The revolution to cloud and Big data. *Chinese Business and Market*, 13(5), pp.41-45.

Liu, Y., 2010. *Organizational Theory Representitive Figures*. Economic Science Press, Beijing.

Luftman, J. & Kempaiah, R., 2007. An Update on Business-IT Alignment: "A Line" Has Been Drawn. *MIS Quarterly*, 6(3), pp.165-73.

Mata, F. J., Fuerst, W. L., & Barney, J. B., 1995. Information technology and sustained competitive advantage: a resource-based analysis. *MIS Quarterly*, 19 (5), pp.487-50.

Melville, N., Kraemer, K. & Gurbaxani, V., 2004. Information technology and organizational performance: An integrative model of IT business value. *MIS Quarterly*, 28(1), pp.283-322.

Morabito, V., 2013. *Managing Digital Information Technology for Value Creation - The SIGMA Approach*. New York.

Nevo, S., 2010. The formation and value of IT-enabled resources: antecedents and consequences of synergistic relationships. *MIS Quarterly*, 34, pp. 163–183.

Nicholsson, N., 1995. *Blackwell Encyclopedic Dictionary of Original Behavior*. Blackwell, Oxford.

Pace, S., 2004. A Ground Theory of the Flow Experiences of Web Users. *International Journal of Human Computer Sciences*, 60(3), pp.327-63.

Pehrsson, P.R., Haytowitz, D. B., Holden, J, M., Perry, C.R.,& Becker, D.G., 1999. USDA's National Food and Nutrient Analysis Program1: Food Sampling. *Journal of Food Composition and Analysis*, 13(4), pp.379-89.

Ridley, D., 2008. *The Literature Review. A step-by-step guide for students.* London: SAGE Publication.

Robson, C., 2011. *Real World Research*. 3rd ed. John Wiley & Sons.

Robson, C, 2002. *Real world research*. Blackwell Publishing.

Savitz, E., 2012. Big data analytics: not just for big business anymore. *Forbes.Com*, p. 4.

Science. 2011. *A special report*: Dealing with data.

Scott, W.R., 2003. *Organization and Organizing: Rational, Natural and Open System*. Pearson Prentice Hall, Columbus, Ohio.

Shuling, Y., 2015. The breif introduce on Big data development in China. *Political Thinking*, 23(4), pp.129-32.

Stephanie, W. & Barbara, W., 2015. Big data: extending the business strategy toolbox. *Journal of Information Technology*, 1(30), pp.60-62.

Steve, L., 2012. The age of Big data. *New York Times*, 11, pp.23-46.

Sturges, J. E. & Hanrahan, K. J., 2004. Comparing telephone and face-to-face qualitative interviewing. *Qualitative Research*, 4(1), pp.107-18.
Uma, S. & Roger, B, 2009. *Research Method for Business - A skill Building Approach*. UK: John Wiley & Sons Ltd.

Taylor, F.W., 1911. *The Principles of Scientific Management*. Norton Library, New York.

Wang, H., 2000. *The Fifty years of Chinese dairy*. Beijing: The Ocean Publisher. pp.352-404.

Weber, M., 1964. *The Theory of Social Economic Organization*. Simon and Schuster Inc., New York.

Xingtang, L., Bingcheng, L., & Guangjun, W., 2008. *Complex System Modeling Theory, Method and Technology*. Science Press, Beijing.

Yang, P.F., 2013. Information productivity and its significant impact on the development prospects. *Edge Revolution 2.0*, 24(3), pp.114-20.

Yin, R.K., 2013. *Case Study Research: Design and Methods*. London: SAGE Publications.

Yue, L., Jia, H., Mingjie, G. & Qing, Y., 2014. An Overview of Big data Industry in China. *Industry Focus*, 11(12), pp.1-10.

Internet

China Mengniu Dairy Company Limited, 2013. *Mengniu Interim Report 2013*. [Online]
Available at:
http://www.mengniu.com.cn/images/pdf/20130926164701001746119_tc.pdf [Accessed 17 April 2016].

CWW Research Report, 2013. *Cwwresearch.com*. [Online]
Available at:
http://www.ccwresearch.com.cn/report_detail.htm?id=131078 [Accessed 03 April 2016].

Dairy Association of China, 2008. The livestock number of cows in China. [Online]
Available at:
http://www.dac.com.cn/search/search.jhtm?gjz=%25E6%2595%25B0%25E6%258D%25AE
[Accessed 10 August 2016].

Guo, J.L., 2015. *The Launching of Yun OS: Expecting to be the 3rd largest mobile phone system in China*. [Online]
Available at:
http://tech.sina.com.cn/mobile/n/n/2015-12-10/doc-ifxmpnqm3039539.shtml [Accessed 12 April 2016].

Liu., Y, T., 2011. *The upgrade of Mengniu's R&D Center*. [Online]
Available at:
http://www.cqn.com.cn/news/zgzlb/diwu/384563.html [Accessed 10 May 2016].

Lohmar, B., F. Gale, F. Tuan, & J. Hansen, 2008. China's ongoing Agricultural Modernization. [Online]
Available at:
http://www.ers.usda.gov/Publications/EIB51[Accessed 11 August 2016].

MathWorks, 2016. *Exporting Data into Database*. [Online]
Available at:
http://images.google.de/imgres?imgurl=http%3A%2F%2Fwww.mathworks.com%2Fcmsimages%2F69760_wl_databasetbfigure4_wl.jpg&imgrefurl=http%3A%2F%2Fwww.mathworks.com%2Fproducts%2Fdatabase%2Ffeatures.html&h=411&w=678&tbnid=Yy2y2pxDrLjgDM%3A&docid=MWnfF2rJ959yKM&ei=a845V7rMCoL7UorNiagE&tbm=isch&client=safari&iact=rc&uact=3&dur=519&page=1&start=0&ndsp=17&ved=0ahUKEwj6rcvv3N7MAhWCvRQKHYpmAkUQMwgkKAIwAg&biw=1314&bih=812 [Accessed 10 May 2016].

Mengniu.com, 2016. *About Mengniu Dairy*. [Online] Available at:
http://www.mengniu.com.cn/about/jtjs/ [Accessed 18 April 2016].

Oracle, 2013. *Bringing the Value of Big data to the Enterprise*. [Online]
Available at:
http://www.oracle.com/us/products/database/big-data-appliance/value-of-big-data-brief-2008771.pdf [Accessed 21 February 2015].

ORACLE, 2015. *Improving Manufacturing Performance with Big data Architect's Guide and Reference Architecture Introduction*. [Online]
Available at:
http://www.oracle.com/us/technologies/big-data/big-data-manufacturing-2511058.pdf
[Accessed 3 March 2016].

Qinbei.com, 2015. *The Big data concerning tacking system on Yili Group*. [Online]
Available at:
http://news.qinbei.com/20150720/1796811.shtml [Accessed 16 April 2016].

Takungpao.com, 2015. *Chinese dairy industry Big data platform launch*. [Online]
Available at:
http://news.takungpao.com/mainland/topnews/2015-10/3224075.html [Accessed 16 March 2016].

Yujie, L. , 2012. *The the analysis of Chinese dairy product in Chinese Market*. [Online]
Available at:
http://202.119.108.161:93/modules/showContent.aspx?title=&Word=&DocGUID=f20ba608daf64483bed40c25b2861a8a [Accessed 16 March 2016]

Yili Group- Our developement , 2014. *Yili offical website*. [Online]
Available at:
http://www.yili.com/about/developPath.html### [Accessed 15 April 2016].

Yiyi Dairy, 2014. *Company information on Yiyi Dairy*. [Online]
Available at:
http://www.yiyimilk.com/aboutyiyi.aspx [Accessed 17 April 2016].

Yiyi Dairy, 2016. *Yiyi Dairy-About Yiyi*. [Online]
Available at:
http://www.yiyimilk.com/active.aspx [Accessed 14 May 2016].

Personal messages

Gao Debu
Company Director, Yili Dairy Group
Phone Interview (2016-04-13)

Luo XX
Representive, Yiyi Dairy Company, sales department
Phone Interview (2016-05-08)

Wu Minglou
CEO, Yiyi Dairy Company
Phone Interview (2016-04-29)

Yao Haichao
Senior R&D Manager, Mengniu Dairy Limited
Phone Interview (2016-04-18)

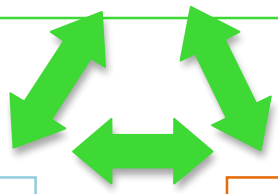
Appendix 1. Interview Guide

Introduction & Key Questions
(Starting with the background info from telephone interview)

- What is your current position and duty in the company?
- What is your educational background?
- How many years have you worked in the company?
- What is your previous working position?

Key Questions

- How do you see digital technology such as Big data can play a big role in your company?
- Could you please describe some advantages if your company plans to use Big data in the future? And what are the main obstacles you will see?



Big data

- Do you know Big data term today?
- How would you define Big data in your company?
- Have you worked with data or statistical related position before?
- Has your company already used Big data? If so, would you please make some using examples in specific areas or products?
- What are the main initiatives when you intend to use Big data?
- Are you able to see the consequences required to leverage Big data?
- What do you think as the greatest benefits/ opportunities concerning the use of Big data? If not, what are the main hindrances?

Business activities & Decision-making

- Would you please give me a brief illustration on how your business activities or possible transactions take place?
- Can you see some great impacts on your business channels (supply, value chain) under the use of Big data in the future? If so, can you explain how these can happen?
- How do you expect to use Big data in order to expand the dairy market? Do you plan to explore some new business areas that are different from dairy business?
- What is your decision attitude towards not using Big data, if so, please give some explanations

Appendix 2. The background of the empirical study

This appendix provides with empirical background information of current situation on Big data development in China. This involves with the developing trend as well as some difficulties in the applicable use of Big data.

Following the global wave of the booming scenario in large data volume, the advent of "Big data" concept is prevailing, and it has been gradually used in various areas of politics, economy, social lives where the embryo of Big data industry ecosystem has constantly been shaped (Steve, 2012). The year of 2011 is regarded as the first year that Big data was introduced as the market tool in China (*ibid*). According to CWW report (2013), the scale of Big data in Chinese market in 2012 was observed to be 470 million RMB, and it was expected to accelerate to 1.12 billion RMB in the next year. Thereafter, the annual growth rate is predicted to be 100% thus by 2016, the market size of Big data will be estimated with the amount to 9.39 billion RMB. Figure 10 (Yue *et al*., 2014) displays the Big data market forecast in China and worldwide. Despite using forecast and different monetary scales, we can observe from this figure that even though the Big data concept is becoming prevailed, the scale of global and Big data market in China is still in the start-up phase. Compared with the global market scale, the Big data in China is considered with the higher growth speed than the global average rate. Followed with such growing speed rate, the Big data market in China will have a huge expansion in the next few years.

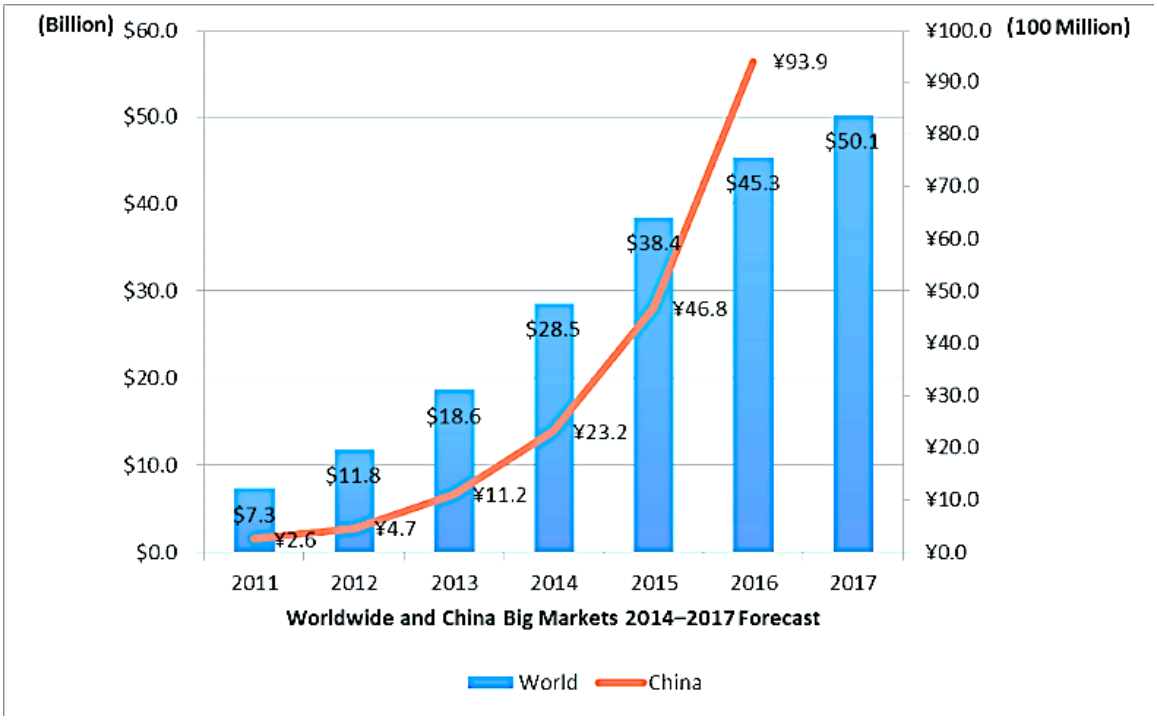


Figure 10. Global and china's Big data market scale 2011-2017. (Yue *et al*., 2014)

This figure also reflects that although companies, organizations, and some scholars have touted the conceptual idea of Big data, the worldwide market including China still lay on the early stage when compared to their entire market scale. However, the prediction to Big data development is potentially huge optimistic. According to Chirkova & Yang (2014), Big data is an unseparated technology, and the growth of Big data highly relies on the Internet medium, particularly the data and application situation of mobile Internet. It cannot imagine

for the usage of Big data without computing of cloud and interconnectedness of all things enabled by computing cloud. Thus, Big data has started to serve business and government issues and thereafter, the surrounding industries such as Internet, finance, telecommunications have paid close attention to Big data development and its growing capacity (*ibid*).

The use of Big data is essential to its intrinsic value and industry flourishing. Brown & Micheal (2011) has presented that "*Big data application works as a 'stepwise' trend: it was driven by the IT industry in the first stage, and then continuously extended to other industries*". Figure 11 demonstrates the Big data market scales from different industry fields in 2013. As it can be realized that the use of Big data in Internet and Government issue have taken about 15% respectively, which are the most extensive usage fields. Finance and telecommunication industries account for 10% individually. However, the energy and manufacture fields seem to have a less usage extent, which only occupy 8.5%.

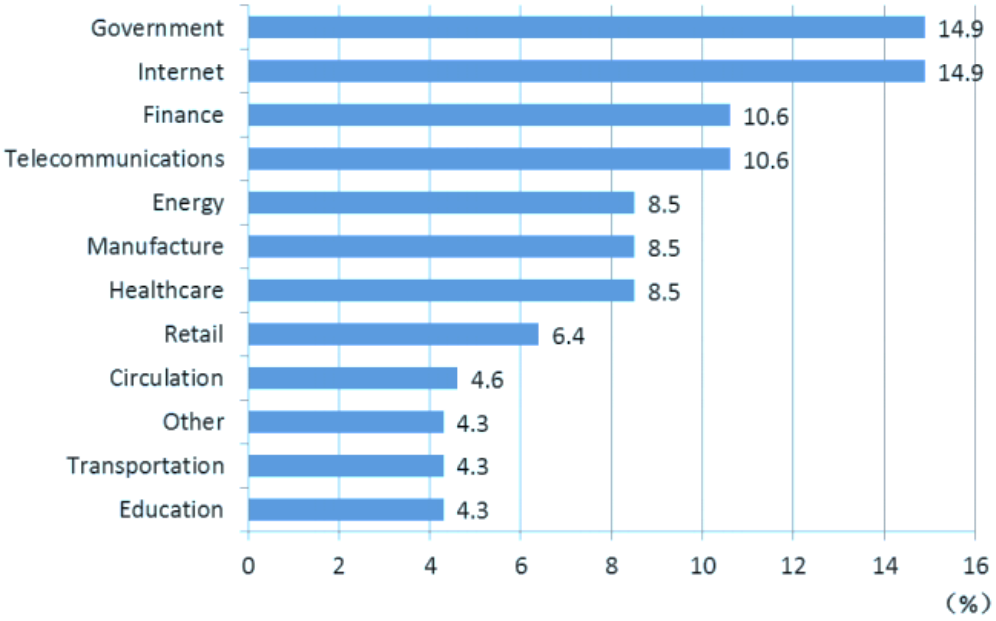


Figure 11. Big data market scale of different types of industries in China. (www, CWW report, 2013).

Regarding Big data application, Yue *et al*. (2014) investigated that the use of Big data can be characterized to three different features in China. To start with, the first feature concerning the Big data application refers to conducting business evaluation and product line optimization based on the data owned by the company themselves; the data are confidential and having the ownership in only one company. The second feature reveals an extended paradigm on Big data application that can permit and share the industry's data license to other industries in vertical fields. Thus, it will be convenient to conduct the extended data analysis in the vertical scope (*ibid*). The last identified feature feature refers to data cross application in two or more industry sectors that more surrounding enterprises will participate in this technological renovation and to ignite their maximum dynamics.

Today, the first feature of Big data application is the most common existing form in China; however, the extended cross features can hardly been detected in among the industry fields. This phenomenon may contribute to fact that Big data companies are always focusing on the individual performance rather than information sharing and collaboration. Meanwhile, some state-owned enterprises have their particular operating system, which considerably prohibit

the information flow and cannot effectively construct a real information ecosystem. Even so, driven by the Big data cross form and advantages of sharing of Internet enterprises, penetration and integration different types of industries in China, the cross form will ultimately become the universal way of using Big data since this new form can create voluminous data, rich information dimension, high value, and leading-edge technologies (Shuling, 2015). Moreover in nowadays, Internet industry is committed to unifying Big data application with finance, public services, healthcare, and other livelihood sectors to perfect a Big data ecosystem. Apart from that, the livelihood group in China is a strong power that can actively enhance the Big data development stage. The expected growing phase from single paradigm to the extended, and finally to the cross form of Big data, rooted in voluminous behavior data of consumers and enterprises, is supposed to stir up a huge industry revolution in China in the next ten years (Yue *et al*., 2014).

Big data challenge for Chinese industries

The Big data market and its relevant industry in China has an extensive prospects for the mature and development; nevertheless, there are still many potential problems occur and the development stage in China is full of challenges which can compare with the scientific revolution in developed countries. The research conducted by Yang (2013) confirms that the application scope of Big data in China is difficult to catch up with the U.S. and other European countires in the short period. On the one hand, Asian countries are not the headstream of Big data industry; meanwhile, the Big data industry in China has already lagged behind the U.S and EU, and the amount of digital data resources are insufficient (*ibid*). On the other hand, the majority of the companies in China believe in their own capability and R&D investment for Big data storage, data processing, and data analysis techniques. This issue leads to chaotic standardization, inaccuracy, low completeness, and a large area of the data source waste in China.

One Chinese scholar Liang (2014) has pointed out that Chinese organizations show the weakness in the fields of databases, data warehousing, and business intellegience. He also maintains that the bigger gap is obvious in our self-control of Big data analysis techniques as well as Big data management in the cross domains. When compared with the U.S, the country have formed a close loop in Big data and are currently generalize this technique as the mainstay of their enterprises. Whereas the Chinese firms are still in the development stage of the close loop. For instance, Google as the most successful search engine worldwide, has accumulated countless data from the Internet through its dominance, and then the company has incrementally build upon the cloud computing technology based on the demand for Big data reserve and analysis (Yue *et al*., 2014). Google can use the more advanced cloud technology to deal with more massive and sophisticated data; thus the company has launched Android and pushing this operating system into the mobile terminals which to exploit a new channel to collect data from users. It is also apparent that a more rich and advanced way of collection Big data can further consolidate its dominance in the technology field, and promote access to Big data (*ibid*). When we compare Big data in Google with Chinese enterprises, the phenomenon reveals that Chinese companies have endeavoured to bridge the technology gap between western countries. Alibaba, which is concerned as the largest e-commerce platform in China, has developed great importance to create a close loop. The company has taken the advantage of its dominant position in e-commerce business, access to the vast amounts of data in order to get benefit for analysing the market competitiveness, customer's buying behavior etc. Ali also imitates Google and also launched its own mobile operating system AliYun OS in 2011 for the purpose of gethering user data in the mobile phone terminal (Guo, 2015).

However, the new system has not obtained a strong reaction due to the immatured core technology and people's cognitive degree. Even though we have to affirm that the development progress of Big data in China has achieved remarkable achievement which reflects on channels where we collect data, platforms we can store data, and teminals we can process data (Shuling, 2015).

To sum up, in the coming event of Big data, China has realize the importance of seizing the new Big data technology wave in order to enhance the production efficiency as well as cross industry collaboration in business, bringing a new vision of economic growth. However, the Big data market scale and its maturity level are still in the infancy stage, and encounter many problems and challenges. The initial will of Big data development in China is to give full play to the role of market mechanism, and allow the government control the various aspects of opening data, formulating laws and regulations, optimizing market conditions, etc. The Big data development in China still has a long way to pursue.