Cape Peninsula
University of Technology

INFORMATION QUALITY PERCEPTIONS IN THE INFORMATION CHAIN OF A RETAIL ORGANISATION

By

JUNIOR MBUYAMBA MUKA

Thesis submitted in fulfilment of the requirements for the degree

Master of Technology: Information Technology

In the Faculty of Informatics and Design

At the Cape Peninsula University of Technology

Supervisor: Prof Retha De La Harpe

Cape Town Campus

January 2015

DECLARATION

I, Junior Mbuyamba Muka, declare that the contents of this thesis represent my own unaided work, and that the thesis has not previously been submitted for academic examination towards any qualification. Furthermore, it represents my own opinions and not necessarily those of the Cape Peninsula University of Technology.

Signed **Date**

ABSTRACT

Dealing with information quality issues remains a serious concern for organisations. With the fact that modern organisations generate much higher volumes of information than they did in the past, with more variety and at a higher velocity, this concern is only expected to increase. Information quality is a complex topic and there is no single view shared by stakeholders on the quality of information within an organisation's value chain and information quality issues experienced by stakeholders are also different. In this study, the topic of information quality is addressed through the following research questions: How do information stakeholders view the quality of information throughout the information value chain of a retail organisation? What kind of information quality issues do business stakeholders experience throughout the information value chain of a retail organisation? And how are stakeholder decisions affected by information quality when producing or consuming information at the different touch points of the information value chain?

An interpretive research approach was chosen for this study to explore this complex topic. The research was conducted in two phases. A literature review in which different concepts related to information quality were investigated was the first phase. In the second phase, a specific business case was investigated before findings from the two phases were analysed and discussed.

In response to the main research questions, the findings of the research indicated that information quality is complex and there is no single view shared by all stakeholders involved in an information value chain.

ACKNOWLEDGEMENTS

I wish to thank:

- My research supervisor Prof Retha De La Harpe for all her patience and guidance throughout this research
- Professor Lorenz Broun from NGU Germany who was always willing to assist me while on a short visit to CPUT
- All my interview respondents without whom I would have had no case study or data to analyse
- Mr Nicholas Challis for the wonderful proof-reading and editing work

Opinions expressed in this thesis and the conclusions arrived at, are those of the author, and are not to be attributed to any of the contributors.

DEDICATION

This thesis is dedicated to:

- Firstly and above all, the almighty God without whom I would never have made it this far. Thank you Lord for your faithfulness.
- To my amazing wife Shingirai Muka for all your support, prayers and encouragements. You are indeed a virtuous wife.
- To my fathers Andre Mbuyamba Muka and Theo Muka, my brother Jean-Paul Muka, my late grand-father George Kitima , and the rest of my big and very supportive family for showing me the way, praying for me and believing in me.
- To all my Cape Town family for standing with me in prayer.

A special dedication goes to my beloved late mother, Georgette Kitima, who would have been extremely proud of this achievement. Thank you for making me realise from a young age that education is important.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DAMA Data management association

DQMDW Data quality meta data warehouse

EIM Enterprise Information Management

ERP Enterprise resource planning

ETL Extract, transform and load

HR Human resources

IaaS Information as a service

IPMap Information product map

IT Information technology

KPI Key performance indicators

PDCA Plan-Do-Check- Act

PSP/IQ Product and service performance model for information quality

SDLC System development life cycle

SOA Service oriented architecture

TDQM Total data quality management

CHAPTER ONE - INTRODUCTION

1.1. Background to the research problem

An information chain is an end-to-end process that starts with original data sources, creates information products and continues through to the use of information in operations, decision making, and planning. It includes supporting technologies such as databases, operational systems, data warehouses and reporting tools (Redman, 2001). Information chains are usually long flows of data that cut across the organisation, department by department. They can also go outside the boundaries of an organisation as data can be supplied from internal and external data sources (De La Harpe, 2008). The University of Alberta (2004) defines an information chain as a production line for converting data into information products that enable decision making. The above definitions suggest that value should be added to data throughout the information chain to produce information products. Hence, the term information value chain will be used in this study to refer to the information chain. As data flow between different points in the information chain, they are processed by different business processes and are subjected to different procedures. In conjunction with business processes and procedures to which data are subjected, there are information stakeholders involved at every touch-point of the information chain. Touch-points refer to stages in the information chain where the interaction between information stakeholders and data takes place. Stakeholders interact with data and the results of the interaction can either improve or reduce the quality of data. Davis (2004) shares this opinion when he stated that whenever data is touched, there is an opportunity to capture, create or propagate poor quality data or to improve the quality of data. It follows from the above that the quality of information products is influenced at every touch-point of the information chain and is highly dependent on stakeholders who interact with data. De la Harpe (2008) is of the opinion that information stakeholders do not only affect the quality of data but can also be affected by the quality of data.

Assumptions: The terms ‘data’ and ‘information’ are often used synonymously (Pipino et al., 2002). It is recognised that these two terms are different, as explained in the ‘data and information’ section below. Where the difference is of little importance, these terms will be used interchangeably. For example, the term ‘information stakeholder’ can be used interchangeably with the term ‘data stakeholder.’ It will also be assumed that the quality of data influences decision making.

1.2. Statement of the research problem

Information stakeholders in different roles interact with data differently. Therefore, information with varying degrees of quality is produced from organisation's information value chains. It is not well understood how different stakeholders view the quality of information and how this influences organisation's decision making.

1.3. Research question, sub-questions and objectives

Table 1.1: Research question, sub-questions & objective

Research question	How do information stakeholders view the quality of information throughout the information value chain of a retail organisation?	
Research sub-questions	**Research method(s)**	**Objectives**
What kind of information quality issues do business stakeholders experience throughout the information value chain of a retail organization?	Interviews Business process analysis Literature analysis Data flow analysis	Establish different types of information quality issues experienced by business stakeholders throughout the information value chain
How are stakeholder decisions affected by information quality when producing or consuming information at the different touch points of the information value chain?	Literature analysis & Document analysis Interviews	Establish whether and how data stakeholders can be affected by information when making decisions

1.4. Current status of research area

This section presents a summary of information quality and discusses fundamental concepts applicable to this research. Before data quality is discussed, it is useful to define the terms: data, information and quality.

1.4.1. Data and information

Although these two terms are often used interchangeably, there is a distinction between them. Data is a representation of things or entities in the real world. It is the raw material from which information is derived. If data is the raw material then information is the finished product (Oz, 2006). Al-Hakim (2007) differentiates these two terms as follows: "*Data are items about things and transactions that are recorded, stored but not organised to convey a specific meaning, while information is data that has been organised in a manner that, it conveys meaning to the recipient.*" These definitions show

that there is a relationship between data and information. They agree with the concept whereby information is a product of an information manufacturing system which will be referred to in this study as the information value chain.

1.4.2. Quality

This section describes the concept of quality in the general market place and examines how this applies to the information product in particular.

A quality product is one that meets its consumers' overall expectations for its intended use. It does not have to possess the best features in every single category and it does not mean luxury either. A product can be perceived as high quality to one customer and poor quality to another (Mehta, 2004).

The below table summarises and describes the fundamental components of product quality in the market place and shows how they apply to the information product.

Table 1.2: Fundamental component of quality applied to information

Component	General application	Application to information
Customer focus	The customer determines quality. Suppliers are successful when they know their customers, understand their needs, understand their market, and consistently strive to meet customer expectations.	The quality of information cannot be improved without an understanding of who information customers are and what their needs are in their information products. There are two main categories of information customers: Internal knowledge workers and external customers.
Process improvement	Improving products and services to meet customer expectations requires improving the processes to remove causes of defects. English (1999) suggests the following actions to improve processes: • Create customer-focused product specifications • Define processes across their value chain • Establish feedback loops throughout the value chain • Encourage team work and	Information quality issues often show the need for a business process reengineering (BPR) or a continuous process improvement (CPI) (English, 2001). • Information specifications must clearly describe the information product required to ensure that it meets the needs of knowledge workers. • Information needs to be defined consistently across its value chain to support

Component	General application	Application to information
	establish a process improvement process.	the needs of all information customers in the value chain. • All players in the information value chain must work as a team. They all need to participate in improving processes.
Scientific methods to process management, control and improvement	Managing, improving and controlling business processes to produce consistent quality results.	Information development processes can be managed, improved, and controlled to produce consistent quality results (English, 2001). The Shewhart cycle, Deming's 14 points of quality and the Juran trilogy are examples of scientific methods to process management that are applicable to information processes

1.4.3. Information quality

Inter-organisational relationships have increased to a great extent in the last decade (Nicolaou, 2009). Due to this, data does not only flow within the boundaries of an organisation but it is also exchanged with other business partners. In addition to flowing within an organisation or between organisations, data is also stored in databases where it is manipulated by data stakeholders (De La Harpe, 2008). Organisations produce much more information than they did in the past. The vast amount of information produced and the manipulations of data by different stakeholders at every touch point of the information chain have increased the focus on information quality. A great deal of research has gone into this field over the years; however it still remains a problem. Waddington (2009) shares this opinion when he states in the information difference survey document that the problem of poor quality information is still very much alive. There are many dimensions to the information quality problem. These dimensions refer to issues that are important to the information consumer. Below are some of the commonly listed information quality dimensions based on Martin (2005):

- Accuracy: Do the data values reflect the properties of the real-world object or event that the data is intended to model?
- Consistency: Are the values of attributes managed in different locations the same?
- Integrity: Are all relationships between data in multiple data stores, tables and files intact?
- Completeness: Do the data contain all relevant information?
- Validity: Do the values held fall between the allowable domains of values established for an attribute?
- Timeliness: The degree to which the latency of data delivery matches the needs of the consuming individuals or processes.
- Relevance: The degree to which the data supports and furthers the goals and objectives of users, processes and the organisation.

The concept of value chain has been used in conjunction with information quality principles to ensure delivery of a quality information product.

1.4.4. Value chain

"A value chain describes the full range of activities which are required to bring a product or service from conception, through the different phases of production, delivery to final consumers, and the final disposal after use" (Klapisnky & Morris, 2001). McCormick and Onjala (2007) define the value chain as a sequence of value adding activities leading to and supporting end users of a particular product. These definitions highlight the value addition that takes place at every stage of the value chain which is an important aspect of this research.

The original value chain model proposed by Porter consists of value activities and margin, with value activities being activities performed by an organisation. There are two types of value activities in Porter's generic value chain: Primary value activities which consist of events that create customer value, and support value activities which facilitate the accomplishment of primary value activities (Dunn et al., 2005).

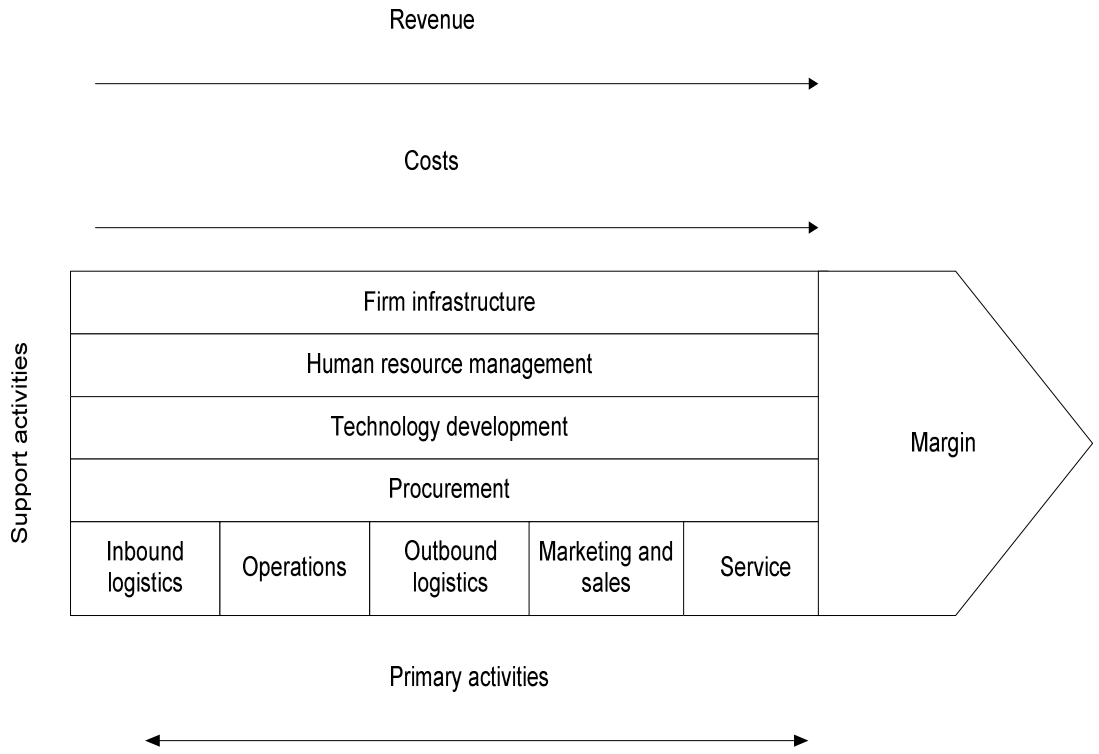

Figure 1.4 Porter's generic value chain, sourced from Dunn et al. (2005)

Value chains can be categorised according to their geographic dimension. They may be national if all processes, from conception to the final disposal after use, take place within national boundaries. They may be international if some processes of the chain spill over and go beyond national borders. They may even be global if different processes take place in different parts of the world (McCormick & Onjala, 2007).

1.4.5. Information value chain

Having established that information is a product (refer to 'data and information section'), the concept of value chain becomes applicable to information processing. English (2001) extends the above definitions of a value chain to define an information value chain as a collection of processes and computer applications that create, update, extract, transform, and present information from its original inception to its final retrieval as well as information presentation to knowledge workers. Laudon and Laudon (2006) agree with English and state that every business has an information value chain in which raw information is acquired and then transformed through various stages that add value to that information. This value addition at the various stages of the information chain is a core aspect of this research. The stages will be referred to as touch-points as it is at

those points that the interaction between information and its various stakeholders happens.

1.4.6. Information stakeholder

"A stakeholder in an organisation is any group or individuals who can affect or are affected by the achievements of the organisation's objectives." (Freeman, 1984; cited in Yläranta 2006) Yläranta (2006) studies many definitions of stakeholder in her study and draws the conclusion that they all have a common message being: "Without the support of stakeholders, an organisation would cease to exist." The same concept can be extended to define an information stakeholder. It is any group or individual who can affect or be affected by the information he / she interacts with. Strong et al. (1997) groups information stakeholders in the following roles: Information producers, information custodians and information consumers. Information producers are people or groups of people who have to deal with the generation of information. Information custodians are people concerned with information storage, maintenance and security. Information consumers are people who use information. This research will look at how these groups of information stakeholders can affect or be affected by the quality of information they interact with.

1.4.7. Decision making

Stakeholders within an organisation use information for different purposes such as supporting their business operations and decisions. With the high volumes of information produced in organisations, it is important for information stakeholders to identify what information of the vast volumes available to them, is of value. Schlogl (2005) defines the purpose of information management as making the right information available, at the right time and at the right place.

In a research aiming to understand information quality in a decision making context, Graefe (2003) used eight criteria which he considered from the two perspectives of information provider and recipient. In his study, he refers to the term information producer as information provider and to information consumer as information recipient. Figure 2 below illustrates these 2 perspectives based on his eight criteria of information quality in the decision process.

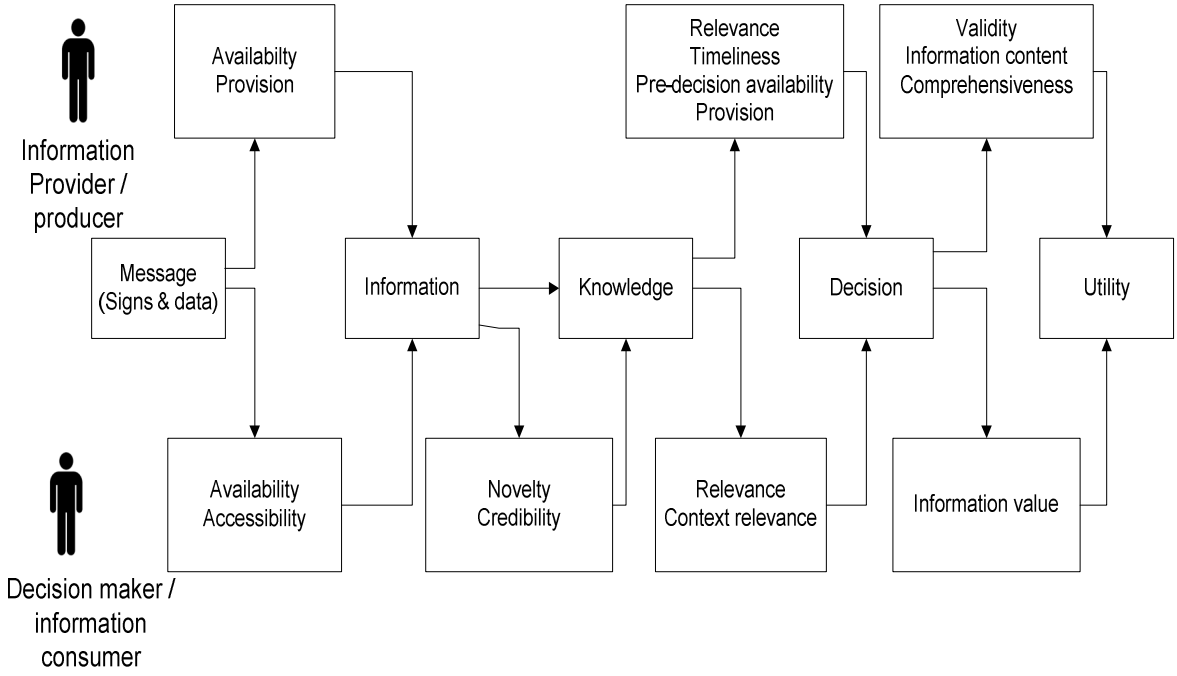

Figure 1.4 Criteria of information quality in decision process, sourced from Graefe (2003)

1.4.8. Conceptual framework

A conceptual framework based on the literature discussed in the previous sections is proposed to indicate the concepts and their relationships that are relevant to this study. The framework is presented in section 2.14.

1.4.9. Research design

The aim of this research is to explore how different business stakeholders view information quality in an organisation's information value chain for a chosen business process, and what kind of information quality issues they experience. The study will make use of a qualitative research paradigm. Niemann (2005) describes qualitative research as any type of research that produces findings not arrived at by statistical procedures or other means of quantification. Empirical research principles will be followed for this study.

The research strategy indicates that a case study is suitable for this investigation. Case studies are mainly exploratory with elements of descriptive research to indicate the flows. A case study was chosen as the preferred strategy because the main research question being posed is a "How" question. Yin (2003) supports this argument when he states the following: "*A case study has a distinct advantage in a situation when a "how" or "why" question is being asked about a contemporary set of events over which the investigator*

has little or no control." A suitable organisation in the retail sector has been identified and has been approached to determine the people who will be involved, the amount of time required from them and the possible expenses.

The unit of analysis is the interaction of stakeholders at each touch-point across the information chain.

1.4.10. Data collection method

A participatory research style will be used to collect the data needed for this study. Participatory research is defined as: "A *method of research where creating a positive social change is the predominant driving force"* (Hughes & Seymour-rolls., 2000). Being an employee of the organisation chosen for the field work, this researcher will use that position to his advantage and gain close involvement to the organisation's utilisation of information and to employees functioning as information stakeholders. Walsham (2006) states that close involvement can have some high benefits because the researcher can be seen by field participants as trying to make a valid contribution.

Interviews and observation will be used to collect practical data. Face-to-face, in-depth, unstructured interviews will be used as the preferred research method. In unstructured, in-depth interviews the researcher aims to collect details, richly textured, person-centred information from one or more interviewees (Niemann, 2005). Boyce and Neale (2006) define an in-depth interview as a qualitative research technique that involves conducting intensive individual interviews with a small number of interviewees to explore their perspective on a specific idea or situation. This researcher will construct a guide of open ended questions as opposed to structured interviews. The main advantage of utilising an open-ended interview format is that it does not force the responses from respondents. Instead respondents are able to answer the questions asked and provide detailed information (Boyce & Neale., 2006).

More data will be collected through the use of an extensive literature review to identify elements that contribute to the data quality issue referred to in the statement of research problem section. The conceptual framework developed above will guide the investigation. The literature will be used to inform the research about issues considered in this study. The findings from the case study will then be compared with the literature.

1.4.11. Theory and data analysis

This research will make use of a theory that considers the socio-technical issues around information quality as found by De la Harpe (2008). In terms of this, the dynamic nature

of how information stakeholders interact with information and view the quality of information throughout the information value chain will be analysed.

1.4.12. Delineation of the research

The business case chosen for this research is one of the biggest retailers in South Africa. This choice was made because their headquarters are located in Cape Town, South Africa where this research is conducted. The case will be looked at strictly from a people, system and information value chain perspective. It is not intended that the outcomes will be representative of the entire retail industry and, therefore, the research outcomes should not be used for comparison or generalisation in that sector.

1.4.13. Contribution of the research

This researcher aims to highlight information quality issues experienced by information stakeholders in an information value chain of a retail organisation. The study will also discuss some of the factors that affect the influence of information stakeholders on information quality.

In addition to the above mentioned feedbacks, the research will include an extensive literature study in the information quality field combined with a case study illustrating the concept of information value chain. The case study results are expected to contribute toward an increased awareness of information as a strategic organisational resource, and a better understanding on the management of information quality throughout the information value chain which could lead to improved decision making. Further research will be conducted to determine whether this can be used in other contexts.

CHAPTER TWO - DATA QUALITY LITERATURE

2.1. Introduction

In this chapter the literature will be reviewed. The literature study is introduced with a discussion on how the terms data and information are viewed in this particular study. The concept of value chains and their relationship with business processes is discussed next. The value chain concept is then applied to information management, resulting in the concept of information value chain. Different interpretations of information quality are discussed before this study's view of information quality is highlighted. Information quality dimensions with related information quality frameworks are discussed. Then follows a discussion on the concept of information stakeholders, information as a product, Information as a service (IaaS) highlighting the relationship between these concepts and the research at hand. Information quality measurement, general information quality management as well as information quality management in the age of big data are also discussed. The last section of this chapter introduces the theoretical conceptual framework developed for this study.

2.2. Data and information

These two terms have been used interchangeably by many authors (Pipino et al., 2002; Chen et al., 2009). However, most authors agree that there is a difference between them, although their interpretation of that difference is not the same. The section below illustrates some of these interpretations.

Data can be viewed as a collection of symbols that signify real world system states and are brought together because they are considered relevant to some purposeful activities. Information on the other hand, is viewed as an objective commodity carried by symbols and relates to who produced it, why and how it was produced and its relationship to the real world state it signifies (Shanks & Corbitt, 1999). In this definition, the information carried by the symbol relates to who produced it, how and why it was produced as opposed to who consumes it. In this study, a more subjective view will be adopted whereby information carried by the symbol / data relates to the consumer who uses it. It is believed that not all information contained in data will be obvious to all consumers who use it as these might have different knowledge, different backgrounds and, therefore, have different interpretations.

Another common definition of these terms is that data are items that are not organised to convey meaning while information is data that has been organised in a manner that, it conveys meaning to the recipient or consumer (Al-Hakim, 2007). In his study, Al-Hakim (2007) clarifies the relationship between data and information. His understanding of that

relationship is consistent with the concept of information as a product of an information manufacturing system, sometimes referred to as an information processing system, which is also how this study views data and information.

Many other authors provide different interpretations of these terms. English (2006) is of the opinion that data can be viewed in many ways. It can be viewed as a representation of things or entities in the real world. For example, consider the sentence: "The author of this thesis, Junior Muka, is a CPUT lecturer." This sentence is data. The entities that it represents are Junior Muka and CPUT. Entity is defined as something that has a separate and distinct existence. Data can also be viewed as facts or attributes. Using the above example, the following can be said: the author's surname is Muka. Surname is an attribute type and "Muka" is the actual value of that attribute type. English (2006) concludes by adding that, data is the raw material from which information may be produced and, information is the meaning of data, facts made understandable. This view is similar to that of Al-Hakim (2007) and is, therefore, aligned to the view adopted in this study.

As stated earlier, the terms ‘data’ and ‘information’ will be used interchangeably in this study, unless specified otherwise. Most authors in the literature suggest that information is data that has been processed, and that data is the raw material that feeds into an information processing system. The information output from one processing system often feeds into another system, and therefore, the information can be data as well. Based on this, it was decided not to differentiate between data and information, unless they are used in a specific context (Davidson et al., 2004). This study refers to the information processing system as ‘the information value chain’.

2.3. The value chain

The concept of value chain is widely used in information management but it originated from business management. It was first described and made popular by Porter (Porter, 1986). Klapinski and Morris (2000) define the value chain as the full range of activities required to bring a product or service from conception, through the different phases of production, delivery to final consumers, and final disposal after use. It is the largest possible process in any organisation and a practical model that allows businesses to analyse activities or processes that create the greatest value as well as those that do not add value. The term ‘value’ shows usefulness or importance. A value chain is, therefore, a model that depicts the increasing importance or value-add, of activities in a process. It is based on a process view of organisations and views a manufacturing (or service) organisation as a system, made up of sub-systems or activities each with inputs, transformation processes and outputs. How the

value chain activities are carried out is directly linked to the organisation's profit margins. (Dunn et al., 2005).

A value chain usually describes a major line of business. Depending on the nature of their business, an organisation can have one or several value chains and each value chain is usually decomposed into several business processes, which in turn, depending on their complexity, can be decomposed into sub-processes, and these are made up of activities. Activities are the smallest sub-process although they can also involve multiple steps. Figure 2.3-1 below depicts the hierarchical representation of business processes.

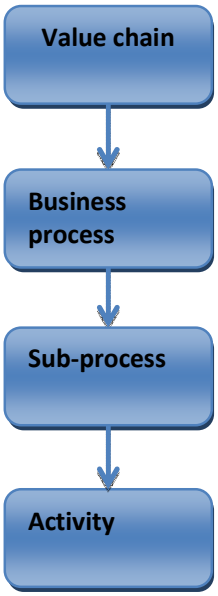

Figure 2.3 Hierarchical representation of business processes – sourced from Ganesan (2013)

Porter's value chain covers two categories of activities, namely; primary activities and support activities (Schmarzo, 2013). The primary activities are:

- Inbound logistics which includes the identification, sourcing and procurement of the raw material needed for the final product or service
- Operations which includes engineering, inventory management, and manufacturing of the final product or service
- Outbound logistics which includes the distribution of the final product and service
- Marketing and sales which includes the marketing, promotions, sales, and customer channel management to get the final product and service to the consumer

- Service which includes the support and maintenance of products and services after they have been delivered to the customer.

The secondary activities are:

- Procurement which includes the acquisition of materials, supplies, and equipment
- Technology development which includes all technologies that support value-creating activities
- Human resource management which includes employee recruitment, training, development, and remuneration
- Infrastructure which includes physical infrastructure such as buildings, offices, and warehouses.

This study looks at value chains from an information management perspective. The study also acknowledges the business management perspective of value chains as activities within a business process are directly related to information flows referred to as information value chains in the study.

2.4. Information value chain

While Porter's value chain focuses on production business processes, many of the concepts he uses can be applied to information management. Schwolow and Jungfalk (2011) demonstrate the application of value chain thinking to information management. They define as primary activities: Information Acquisition, Information Processing and Information Distribution. On the other hand, they identify knowledge management, information governance, human resource management and IT infrastructure as support activities. They describe their model as a tool for information managers to assess the state of information management in their organisations.

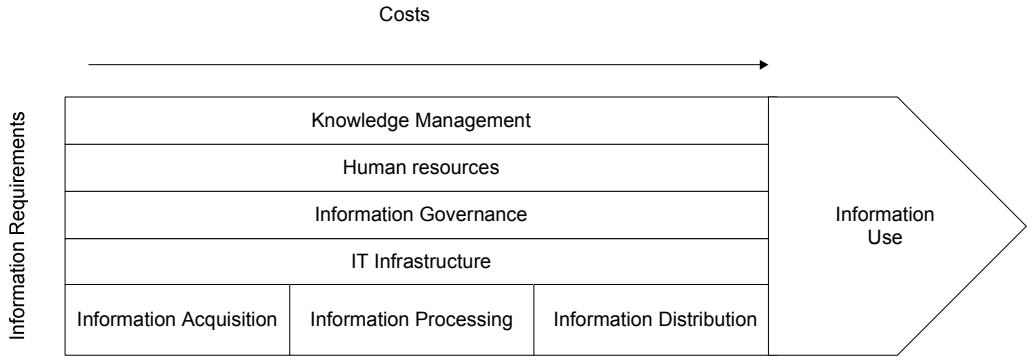

Figure 2.4 Information Value Chain management - sourced from (Schwolow & Jungfalk., 2011)

Information requirements: This is the first activity in the information value chain and it includes a detailed description of the organisation's information needs. Based on its needs, the organisation defines the information products or information resources required. Schwolow and Jungfalk (2011) add that the organisation must be able to measure the cost of these information resources.

As mentioned above, Schwolow and Jungfalk (2011) distinguish between the primary activities and support activities, just like Porter's value chain model. The primary activities of the information value chain are:

- Information acquisition which consists of internal and external environmental scanning processes and adds value by identifying and acquiring the raw material or data that goes into the creation of information products during information processing or acquiring the finished information products

- Information processing which refines the raw material or data acquired during environmental scanning and it adds value through editing, preparation and representation of information products for specific purposes

- Information distribution which involves activities needed to distribute and share information products and it adds value through utilisation of economies of scales, lowering the cost per information product.

The supporting activities are:

- Knowledge management which covers all processes with the objective of exploiting, which includes identifying, representing or distributing, information that exists in the form of 'information-as-knowledge.' It is used to create, store, share and use organisational knowledge and experience. It adds value by turning tacit or hard-to-measure information into manageable information products. Schwolow and Jungfalk (2011) go further by arguing that knowledge management can facilitate organisational learning

- Human resource (HR) management which covers HR-related issues such as recruitment, remuneration and training. It adds value through the recruitment of people who understand the importance of information

- Information governance which provides policies, standards and guidelines for handling information resources in the organisation, throughout the entire information

life-cycle, from acquisition, processing, distribution and storage, to maintenance and disposal of information resources

- IT infrastructure which consists of all back-end and front-end systems in the organisation and adds value through automation of business processes and increasing the speed and quality of information processing.

Understanding the information value chain related to business processes will be of great benefit to information management for any organisation (Schwolow & Jungfalk., 2011). By taking this concept and narrowing it down to information quality management, which according to DAMA (2010) is a function of information management, the following can be said: A good understanding of information value chain related to business processes will be of great benefit to information quality for any organisation. In most organisations, information systems no longer work in isolation. Information flows from upstream systems to many downstream systems. Jones (2008), in his attempt to list essential skills for information quality management, mentions that an understanding of how information flows around your business to deliver value to your internal and external customers is critical to effective information quality management.

As mentioned earlier and confirmed by Berner (2012), an information value chain focuses on the set of activities that add value to information as it flows throughout the organisation The aim is to increase the usefulness of information to the users, enabling them to make better decisions. This focus on increasing the usefulness of information is directly related to this study's understanding of information quality and, therefore, makes the information value chain an important concept to grasp for proper information quality management.

2.5. Quality of information

The previous section looked at the concept of value chain and its application to information management in the form of information value chains. This section looks at the different interpretations of information quality and highlights the view that will be adopted in this study.

Quality information is generally defined as information that is fit-for-purpose or fit-for-use (Neely, 2005; Knight, 2011; Jiang et al., 2008); it is about whether information meets implicit and explicit expectations of people who will use it. These expectations can be very complex as they are based not only on what the information is meant to represent, but also on why the user needs the information and how he/she intends to use it (Sebastian-Coleman, 2013). This view of quality suggests that information considered of high quality for one use may not

be good enough for another use. The 'fit-for-use' paradigm to information quality embraced by many researchers, as shown above, also implies that information quality cannot be defined and assessed outside the reasons for which it exists (Knight, 2011).

However, it is worth mentioning that not all information quality researchers agree with the fit-for-use paradigm. For example, Shanks and Corbitt (1999) are of the opinion that information quality should be assessed in the context of its generation as opposed to its use. Herzog et al. (2007) highlight that in many settings, especially for intermediate products, data quality is often viewed as conformance to standards set. This study, however, leans more toward Knight's understanding of quality as 'fitness-for-use' because this paradigm recognises the widely accepted multi-dimensional perception of information quality (Karr et al., 2001).

Jiang et al. (2008) propose an approach to understand and define information quality attributes. In their study, they agree that information quality is not only multi-dimensional but it is also a hierarchical concept. An example of the hierarchical view of information quality is that credibility can be considered a sub-dimension or sub-attribute of the believability dimension. This hierarchical view of data quality, although agreed with, will not be explored in this study. The general consensus in literature is that information quality should be taken as encompassing multiple dimensions. It is for that reason that this study only focuses on the multi-dimensional aspect and not the hierarchical aspect.

2.6. Information quality dimensions

Several authors in the information quality literature have attempted to classify and define the various aspects of information quality often referred to as information quality dimensions. Strong et al., (1997) define an information quality dimension as "a set of information quality attributes that represent a single aspect of information quality." Loshin (2011) looks at a dimension as a perspective or a frame of reference for measurement. Sebastian-Coleman (2013) agrees with Loshin (2011) when he states that information quality thinkers have used the word dimension to identify aspects of information that can be measured and through which quality can be quantified. As discussed in the previous section, information quality is also viewed as a hierarchical concept. As a result of that shared understanding, there have been several attempts to group or classify information quality dimensions (Strong et al., 1997; Dedeke, 2000; Naumann & Rolker, 2000; Kahn et al., 2002). Not all of these classifications have been accepted and adopted by all information quality researchers. Some authors have been criticised of being ambiguous and inconsistent in their classifications and definitions of information quality dimensions (Jiang et al., 2008). These critics have led to

many others attempting to re-classify and re-define information quality dimensions. Strong et al., (1997) is probably still the most cited as far as information quality classifications or categories, and dimensions within those categories are concerned. This study does not attempt to re-classify Information quality dimensions but uses existing classifications or frameworks to understand Information quality issues for the chosen business case. The following section looks at some of the information quality frameworks found in literature and highlights their related dimensions.

2.7. Information quality frameworks

As mentioned in the previous sections, there have been many attempts by information quality researchers to classify and define information quality dimensions. This sections looks at classifications or information quality frameworks that have been proposed by information quality authors, drawing from a study by Knight (2011) who assesses a total of twenty one information quality frameworks, and finds that they share a number of characteristics regarding their classifications and descriptions of information quality dimensions, with some of the frameworks having several dimensions in common.

In the categorical information quality framework, Strong et al., (1997) insist on the importance of information quality categories and dimensions. They argue that the conventional intrinsic view of information quality, which is independent of the context in which information was produced and consumed and only focuses on errors in stored data, is not sufficient to solve information quality issues in organisations. As a result, Strong et al., (1997) broaden the intrinsic view of information quality to include data in production and utilisation process. They categorised dimensions of information quality into Intrinsic, contextual, representational, accessibility. Table 2.1 depicts these categories and their dimensions.

Table 2.1: Categorical information quality framework

Information Quality Category	Information Quality Dimensions
Intrinsic Information Quality	Accuracy
	Objectivity
	Believability
	Reputation
Accessibility / Interactional Information Quality	Accessibility
	Access security
Contextual Information Quality	Relevancy
	Value-added
	Timeliness
	Completeness
	Amount of Data
Representational Information Quality	Interpretability
	Ease of understanding
	Concise representation
	Consistent representation

The Semiotic information quality framework was developed by Shanks and Corbitt (1999) based on context from Semiotic theory, which uses symbols to transfer knowledge. In their framework, four levels of information quality are distinguished: Syntactic, semantic, pragmatic and social. Information

Table 2.2: Semiotic information quality framework

Quality category	Category Description	Information quality Dimensions
Syntactic information quality	Concerned with the structure of information and information conformance to meta data (i.e. database rules)	Accuracy
		Believability
		Objectivity
		Reputation of the source
Semantic information quality	Concerned with the meaning of information as per the user	Value-Added
		Relevancy
		Timeliness
		Completeness
		Amount of data
Pragmatic information quality	Concerned with the importance of information for use	Interpretability
		Ease of understanding
		Representational consistency
		Concise Representation
Social information quality	Concerned with the shared understanding of information by various social groups	Accessibility
		Ease of Operations
		Security

A conceptual framework for measuring information quality is proposed by Dedeke (2000). In this framework, Dedeke identifies five categories of quality, and three of these categories, namely; accessibility, contextual and representational, are aligned with the categories

mentioned by Strong et al., (1997). To develop this framework, Dedeke start by identifying data, interfaces, work, hardware and software as some of the key components of an information system. He then goes on to identify quality categories based on the relationships between these components.

Table 2.3: Conceptual framework for measuring information quality

Information Quality category	Information quality dimensions
Ergonomic Quality	Ease of Navigation, Confortability, Learnability, Visual signals, Audio signals
Accessibility Quality	Technical access, System availability, Technical security, Information accessibility, Information sharing, Information convertibility
Transactional Quality	Controllability, Error tolerance, Adaptability, System feedback, Efficiency, Responsiveness
Contextual Quality	Value added, Relevancy, Timeliness, Completeness, Appropriate information
Representation Quality	Interpretability, Consistency, Conciseness, Structure, Readability, Contrast

In the classification of information quality metadata criteria suggested by Naumann and Rolker (2000), twenty two information quality criteria are summarised into three assessment classes.

Table 2.4: Classification of information quality metadata criteria

	Dimensions
Subject Criteria	Believability
	Concise representation
	Interpretability
	Relevancy
	Reputation
	Understandability
	Value-Added
Object Criteria	Completeness
	Customer Support
	Documentation
	Objectivity
	Price
	Reliability
	Security
	Timeliness
	Verifiability

	Dimensions
Process Criteria	Accuracy
	Amount of Data
	Availability
	Consistent representation
	Latency
	Response time

The product and service performance model for information quality (PSP/IQ) proposed by Kahn et al. (2002) focuses on the product and service characteristics of information, specifically the 'conformance to specification' definition of product and 'meeting or exceeding customer expectation' view of service.

Table 2.5: Product and service performance model for information quality

Quality type	Classification	Dimension
Product Quality	Sound Information	Free-of-Error
		Concise representation
		Completeness
		Consistent representation
	Useful information	Appropriate Amount
		Relevancy
		Understandability
		Interpretability
		Objectivity
Service Quality	Dependable Information	Timeliness
		Security
	Useable Information	Believability
		Accessibility
		Ease of Manipulation
		Reputation
		Value-Added

The conceptual framework for information quality in a website context by Eppler and Muenzenmayer (2002) consists of four manifestations, two quality categories and sixteen quality dimensions.

Table 2.6: Conceptual framework for information quality in a website

Information Quality types	Information quality categories	Dimensions
Content Quality	Relevant Information	Comprehensive
		Accurate
		Clear
		Applicable
	Sound Information	Concise
		Consistent
		Correct
		Current
Media Quality	Optimized Process	Convenient
		Timely
		Traceable
		Interactive
	Reliable Infrastructure	Accessible
		Secure
		Maintainable
		Fast

Although different in their approaches, it can be easily noticed that many of these frameworks have common information quality dimensions. For example, accuracy, accessibility, completeness, timeliness, and consistent representation are all dimensions cited in more than one information quality frameworks discussed above.

In order to have a good understanding of information quality, it is not enough to look at all classifications or information quality frameworks. It is also very useful to look at different stakeholders involved in the generation, management and more importantly the use of information (Knight & Burn., 2005). The next section focuses on stakeholders involved in the information value chain.

2.8. Information stakeholders

The most accepted definition of a stakeholder in an organisation is any group or individual who can affect or is affected by the achievements of the organisation's objectives (Freeman, 1984 cited in Bryson 2004; Yläranta 2006; Jeffery 2009). Jeffery (2009) adds that stakeholders have an interest in the operations and decisions of the organisation and that they supply resources that are critical to the success of the enterprise. Yläranta (2006) studies many definitions of stakeholder in her study and draws the conclusion that they all

have a common message being: “Without the support of stakeholders, an organisation would cease to exist.” This study applies the same concept to information stakeholders.

An information stakeholder is an individual or a group that can be affected by or can affect a situation with information required to perform specific activities. Information stakeholders are usually humans; however, it is also possible for an organisation or a department within an organisation to be considered an information stakeholder (De la Harpe, 2008). Examples of information stakeholders in a retail organisation are: a merchandise planning manager, a merchandise buyer, a till operator, a supply chain manager, a store manager, the marketing department, the information technology department etc. All these information stakeholders have an interest in how information products are designed, and how they can use them.

Information stakeholders interact with information in the roles of information producer, consumer, custodian and manager (De la Harpe, 2008; Shanks & Corbitt, 1999). Often, information stakeholders in different roles have different priorities on information quality dimensions. In order to meet the information quality requirements of all stakeholders’ roles, it is essential to understand which quality dimensions are important to them. The section below is a summary of a study conducted by Giannoccaro et al. (1999) in which the differences between the information stakeholders’ roles and information quality dimensions are highlighted that they consider to be a high priority.

2.8.1. Information producers

Information producers are people who create or collect information (De la Harpe, 2008). Examples of information producer in a retail organisation include a till operator. The creation and collection of information is usually done within a specific context. As a result, they usually perceive the contextual and representational information quality categories to be more important. Particularly, the ‘amount of information’ and ‘ease of understanding’ dimensions within those categories (Strong et al., 1997).

2.8.2. Information custodians

Information custodians are people who design; develop; and maintain systems that store information (De la Harpe, 2008). An example of information custodians includes the Information technology department. They are responsible for providing and looking after resources for storing, processing and accessing the information. Based on this, it is considered that they are more interested in the intrinsic and accessibility category, in particular the ‘security’ and ‘accuracy’ dimensions (Strong et al., 1997).

2.8.3. Information consumers

Information consumers are those who use the information as part of their jobs (De la Harpe, 2008). An example of information consumers in a retail organisation includes merchandise planners. For decision making purposes, consumers need information to be delivered to them in a timely, complete and accessible manner. As a result, they are considered likely to perceive the accessibility and contextual category as most important. In particular, the accessibility, timeliness and completeness dimension (Strong et al., 1997).

2.8.4. Information managers

Information managers are people responsible for managing the entire information value chain, including the accuracy, security and representation of the information (De la Harpe, 2008). An example of information managers is the Enterprise Information management team (EIM). They are considered likely to perceive the intrinsic, representational and accessibility categories as most important, with more emphasis on the accuracy, interpretability and security dimensions (Giannoccaro et al., 1999).

The diagram below shows the proposed association between stakeholder roles and information quality categories using the categorical information quality framework.

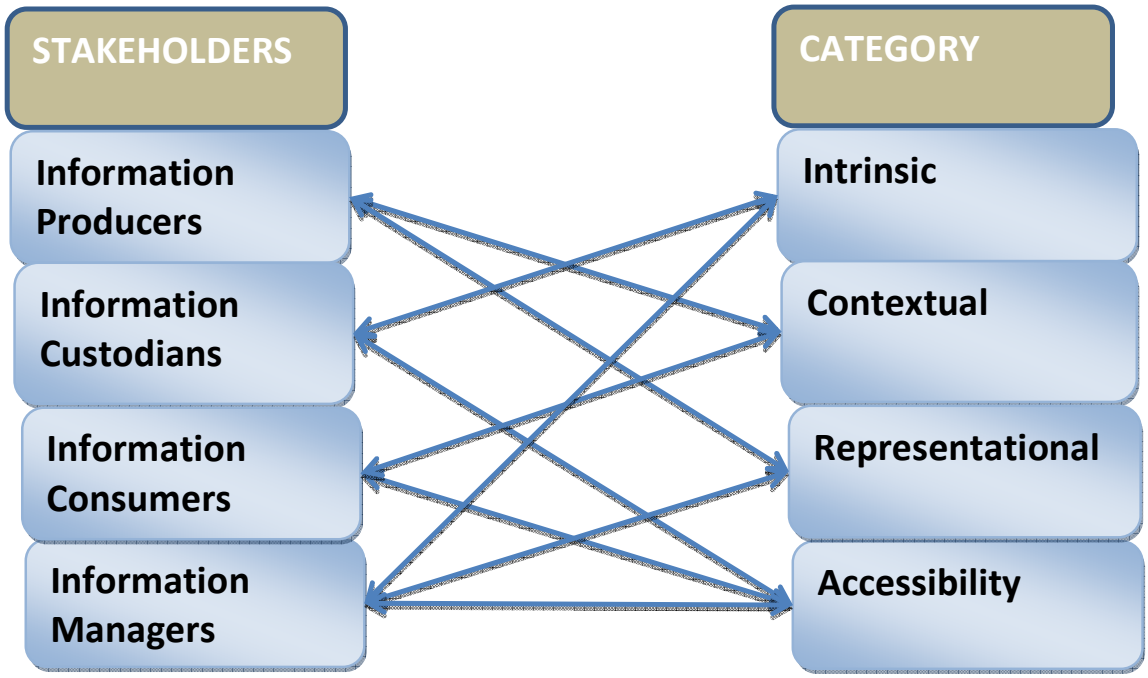

Figure 2.8 Associations between stakeholder types and information quality categories, based on Strong et al. (1997) and Giannoccaro et al. (1999)

Not every author in the literature is in complete agreement with Giannoccaro et al. (1999). In their study, Lee and Strong (2003) agree that there is a relationship between information stakeholders and information quality dimensions. They state that there is a relationship between knowledge and information quality and that this relationship differs by information roles. However, their associations between information quality stakeholders and information quality dimensions are not the same as those established by Giannoccaro et al. (1999). They associate information producers' knowledge with accuracy and completeness dimensions of information quality. Information custodians' knowledge associated with completeness, accessibility, and timeliness dimensions of information quality. Information consumers' knowledge is associated with relevancy of the information. They focus on five dimensions of information quality: accessibility, relevancy, timeliness, completeness and accuracy. As a concluding remark, they state that knowledge of information production processes or information chain processes indicates the competency for producing high quality information. As discussed in section 2.3, the view in Lee and Strong (2003) concluding remark is shared in this study.

At least three of the four information roles can be found on every touch-points of an information chain. For example, some touch-points do not have information producers because the information might have been collected in an up-stream system and fed through to the downstream system. A lack of understanding of the information value chain and what information roles perceive to be high quality information will often lead to serious dissatisfactions and distrust in the information. To illustrate this point, the following can be said: Whenever information custodians and information managers do not understand their information value chain, information consumers will most likely lose trust in the quality of information produced from that information chain. (Wang et al., 2003)

2.9. Information as a product

A product is a tangible item that can be bought, owned, stored and even sold again at a later stage. It can also be used and re-used in the future. Viewing information as a product, takes into consideration the manufacturing process for which the end-product is information stored in a database (Kahn et al., 2002). This end product of information manufacturing systems referred to as information value chains in this article, can be bought, owned, stored, used and reused like any other product.

To achieve the delivery of high quality information, many authors and information quality professionals agree that information must be managed as a product (Wang et al., 2003). This suggests that organisations need to focus on delivering information as a product to its consumers. Dekeke (2005) shares this opinion when he states that experts in the field of information quality have realised the need for a paradigm shift in order to resolve information quality issues in firms. That paradigm shift consists of treating information as a product and aims at ensuring the delivery of high quality information product. A quality information product conforms to the specifications. That is, it meets standards of accuracy, completeness, and freedom-from-error. It must also meet consumers' expectations by being useful and relevant to their needs (Kahn et al., 2002: 186). All this will not be successful unless the entire information chain is well managed and understood as stated in the previous section.

Wang et al.,(2003) propose four principles for treating information as a product based on many different fields of studies and organisational settings. The section below discusses the four principles:

- **Understand information consumers' needs**

  In addition to leveraging on the well-established system analysis and database management processes and techniques, information systems and information quality professionals, as well as researchers, also need to capture, store, analyse and report on information quality meta data. This means capturing data about the quality of information and the resulting activities. This meta data includes consumers' expectations of information quality based on the different information quality dimensions discussed earlier on in this chapter (Loshin, 2011), bearing in mind that there are many data consumers that use the final information products delivered for decision making. In support of this concept, Maydanchik (2007) uses the term 'data quality meta data warehouse (DQMDW)' to describe the collection of tools and processes for collecting, organising and analysing the various types of meta data that is utilised or created while managing information quality, which he also describes as critical.

- **Manage information as the product of a clearly defined information value chain**

  The information value chain concept has been discussed earlier in this chapter. Wang et al.,(2003) refer to it as the information product map (IPMap) and acknowledge that it is an essential concept for managing information quality. In their

study, they find that not many organisations have clearly defined and documented information value chains and that contributes to the poor quality of information products delivered. The importance of having well-defined information value chains is illustrated by the fact that quality of information deteriorates when moving across functional areas or organisational boundaries, and these movements should be represented in the information value chain.

- **Manage the life cycle of the information product**

  To manage the life cycle of the information product, Wang et al., (2003) recommend the use of information quality software tools and techniques to assess the quality of organisational information. They distinguish between software tools for subjective assessment and tools for objective assessment. Subjective assessment refers to information stakeholders' evaluation of organisational information quality based on subjective information quality dimensions while objective assessment means assessing information quality using integrity rules.

- **Appoint information product managers**

  This refers to the need to establish an entity that is responsible of overseeing the information product life cycle across functional areas of the organisation or across organisational boundaries. This role is responsible for tasks such as, raising awareness on the importance of managing information as a product, and establishing the practice of managing information as a product in the organization. Wang et al., (2003) insist that this management component is crucial for the success of information quality management in an organisation. Without appointing people who have the responsibility, authority and correct skill to manage the information as a product, the final information product delivered will not meet expectations of information consumers.

To improve the management of information quality, information can also be viewed as a service in addition to being treated as a product. The next section looks at information as a service.

2.10. Information as a service (IaaS)

A service is defined by the Oxford dictionary as "assistance or advice given to customers during and after the sale of goods." It is produced and consumed simultaneously. As opposed to a product, a service cannot be stored. Although a product is very different to a

service, it is not sufficient to only treat information as a product. It must also be viewed as a service. Failing to offer information as a service, misses some aspects of high importance to the consumer (Kahn et al., 2002: 186; Shanks & Corbitt 1999) as the delivery of information to its consumers has characteristics of a service. It often involves personal interactions between the information delivery staff, often part of the information technology (IT) department, and the information consumers.

For many years, information consumers in organisations have struggled to gain access to correct information, at the right time and in the right format. Treating information as a service allows this difficult goal to be achieved (Reeve, 2013). That is because it allows the distribution of information across the enterprise as a shared service as well as standardised access to information. It enables the 'information-on-demand' paradigm by allowing services to access the information when needed. In turn, the distribution of information as shared services across the enterprise allows the organisation to cut down on the time and cost involved in physically extracting information from one system, transforming the data into the format needed by the consumer, and loading the data into the reporting system (ETL), therefore, making the information available to the consumer much sooner. In other words, the information-on-demand paradigm simplifies the organisation's information value chain. The open data center alliance, inc. (2013), in their first review of the master usage model for information as a service, lists the following as benefits of information as a service, which are aligned to the ones mentioned above:

- Dynamic ability to acquire information through an on-going coordination of information delivery from multiple data sources and in multiple formats
- Reduction of cost, time and complexity of sharing information
- Standard methods that allow information producer, custodians, consumers and managers to interact easily, consistently and efficiently
- Separation between information consumers and the information that they consume.

The concept of information as a service is part of a design philosophy in which software is designed and built in pieces that provide well-defined services when requested. This design philosophy is referred to as service oriented architecture (SOA) (Reeve, 2013).

Similarly to the view of information as a product, an information service has to conform to specifications in order to be deemed of good quality. In this case, it means that the information delivered to a consumer must be dependable. It must be made available to the consumer always on time and in a secure manner. Furthermore, an information service must meet the consumer's expectations by being easy to obtain, manipulable and adding value to their tasks (Kahn et al., 2002). As a service, the quality of information also focuses on the

consumer's response to their task based interaction with the information system (Shanks, 2004).

2.11. Information quality measurement

According to Loshin (2011), a dimension for information quality describes a context and a frame of reference for measurement, a means through which information quality can be measured. Therefore, having discussed the concept of information quality dimensions earlier in this chapter evokes thoughts of measurement. Information quality is generally viewed as a subjective topic such that it cannot be assessed or measured independently of the users (De La Harpe, 2008; Caballero et al., 2007). Loshin (2011) agrees with this when he states that the fitness of information for any specific purpose is dependent on the needs of the user. At every touch-point in an information chain, information is used by different roles having different requirements on it for the tasks assigned to those roles. For each role, information has different values at different moments. Therefore, performed measurement may have a different meaning. For this reason, this study also supports the opinion that the quality of information in an information value chain cannot be assessed independently of the people using the information.

However, some objectives methods for assessing information quality have also been developed. In her study, De La Harpe (2008) explores many frameworks and guidelines for assessing information quality and indicates some problems associated with measuring information quality. Some frameworks were subject-area specific, such as data warehousing or enterprise resource planning (ERP) specific. Others were more generic, but all were said to have short-comings in terms of how true their information quality scores are.

In addition to the subjective versus objective methods of assessing information quality, Caballero et al. (2007) distinguishes between a static and a dynamic way to measure the quality of information. The dynamic way is a set of static measures performed at different points of the information value chain. This could mean taking different snapshots for tracking the different information quality values through the information value chain. This dynamic method of measuring information quality, which sounds very objective, could be made subjective by involving users in all static measures performed at every touch point. Therefore, it is also aligned to this study's interpretation of information quality measurement.

The ability to assess and measure the quality of information is important in Information quality management. Organisations need to measure the quality of information throughout their information value chains to ensure that the resulting information product is of a good

quality. If you do not measure quality at different points in the chain, you do not know whether it is improving or degrading, and that would be contradictory to what information quality management is all about. One of the most quoted management mantra says that you cannot manage what you don't measure. This mantra might have been an inspiration from the famous British mathematical physicist, William Thomson also known as Lord Kelvin, when he stated: "If you cannot measure it, you cannot improve it." Nevertheless, the concept is very much applicable to information quality management (Sebastian-Coleman, 2013).

2.12. Information quality management

As a function of the broader information management, information quality management emphasises managing information the way one would manage any other enterprise asset. It's the "planning, implementation and control activities that apply quality management techniques to measure, assess, improve and ensure the fitness of information for use" (DAMA, 2010). It's not only about fixing the organisation's information related issues, but it's rather about managing the lifecycle of information creation, transformation, and transmission to ensure that the resulting information meets the needs of all information consumers within the organisation. In other words, information quality management is about managing information throughout the information value chain. The DAMA's view of information quality management sounds somewhat misaligned to the English (2010**)** definition, when he states that information quality is an extension of quality management and not an extension of information management, therefore, proven quality management principles, processes and techniques that have been successfully implemented in other industries such as product manufacturing, should also be applied to information products. Although DAMA's view of information quality as a function of information management appears to be contradictory to that of English (2010), some of the quality principles on which they are based are shared.
The table below looks at some these quality principles and techniques:

Table 2.7: Seven information quality principles; adapted from English (2010) and Friedman (2014)

Principle	Explanation
Information quality is a business problem and not just a system problem	The goal of information quality management is not just to improve data stored in Data warehouses and source databases. It is to increase business effectiveness by eliminating the cost of non-quality information and by increasing the value of high quality information assets. Information quality solutions are business solutions that encompass Business processes, Applications, databases and people.
Focus on people, not just on Information	When starting an Information quality programme, it is very tempting to only focus on the information. Organizations need information because it is required by information consumers to make business decisions and perform business processes. English (2010) quotes W. Edwards Deming's 14 points of quality which describes an obligation to a customer that never ceases because the consumer is the most important part of the production line. In information quality management, this means that the obligation to the information consumer never ceases as he is the most important part in the information value chain.
Implement information quality management processes, not just information quality software	Organizations often introduce information quality software to solve information quality problems by analysing or cleansing the information. However, failure to identify the root cause of these information quality issues often causes these solutions to fail. Although information quality software play an important role in information quality management, it is a mistake to believe that the software will solve all information quality problems. Organizations need to understand the business problem that they are solving, implement quality management processes, then select the appropriate information quality software.
Emphasize process improvement and preventative maintenance, not just corrective maintenance	The highest payoff for quality information is achieved when we designed quality in the processes that produce information. English (2010) refers to W. Edwards Deming once more when he states "Cease dependence on inspection to achieve quality. Eliminate the need for inspection on a mass basis by building quality into the product in the first place." This doesn't mean that organizations should not conduct clean-ups or corrective maintenance. It means that these clean-ups should be once-off activities and that in parallel to the clean-ups, organizations must conduct a Plan-Do-Check- Act (PDCA) process improvement to identify the root cause of the non-quality and implement improvements to prevent recurrence.
Improve processes at the source, not just in downstream systems or business areas	Information quality problems should be resolved at the source where the information comes from and that quality must be maintained throughout the information value chain wherever value is added to information.
Provide training to information stakeholders	Information Quality training must be provided to all information stakeholders involved in the information value chain. For example, Information managers must understand the information value chain of the organization, including all information producers and information consumers involved. Information producers must understand the meaning of data, the valid values, the business rules, as well as how to produce information that meet the needs of information consumers.

As highlighted in principles above, the successful management of information quality involves people, process and technology. Several frameworks have been developed to help organisations manage and improve the quality of their information. One of the most popular frameworks is the Total data quality management (TDQM) developed by Dr Stuart Madnick who was inspired by the success of total quality management in manufacturing (Talburt, 2011).

In this study, we shall look at a framework developed by Loshin (2011), which he calls the virtuous cycle for information quality. This is a process to help organisation improve the quality of their information assets. The framework strongly supports the view that that the quality of information cannot be assessed in isolation from the information consumers, and it also recognises the multi-dimensional aspect of information quality. It is for these reasons that it has been chosen for this study.

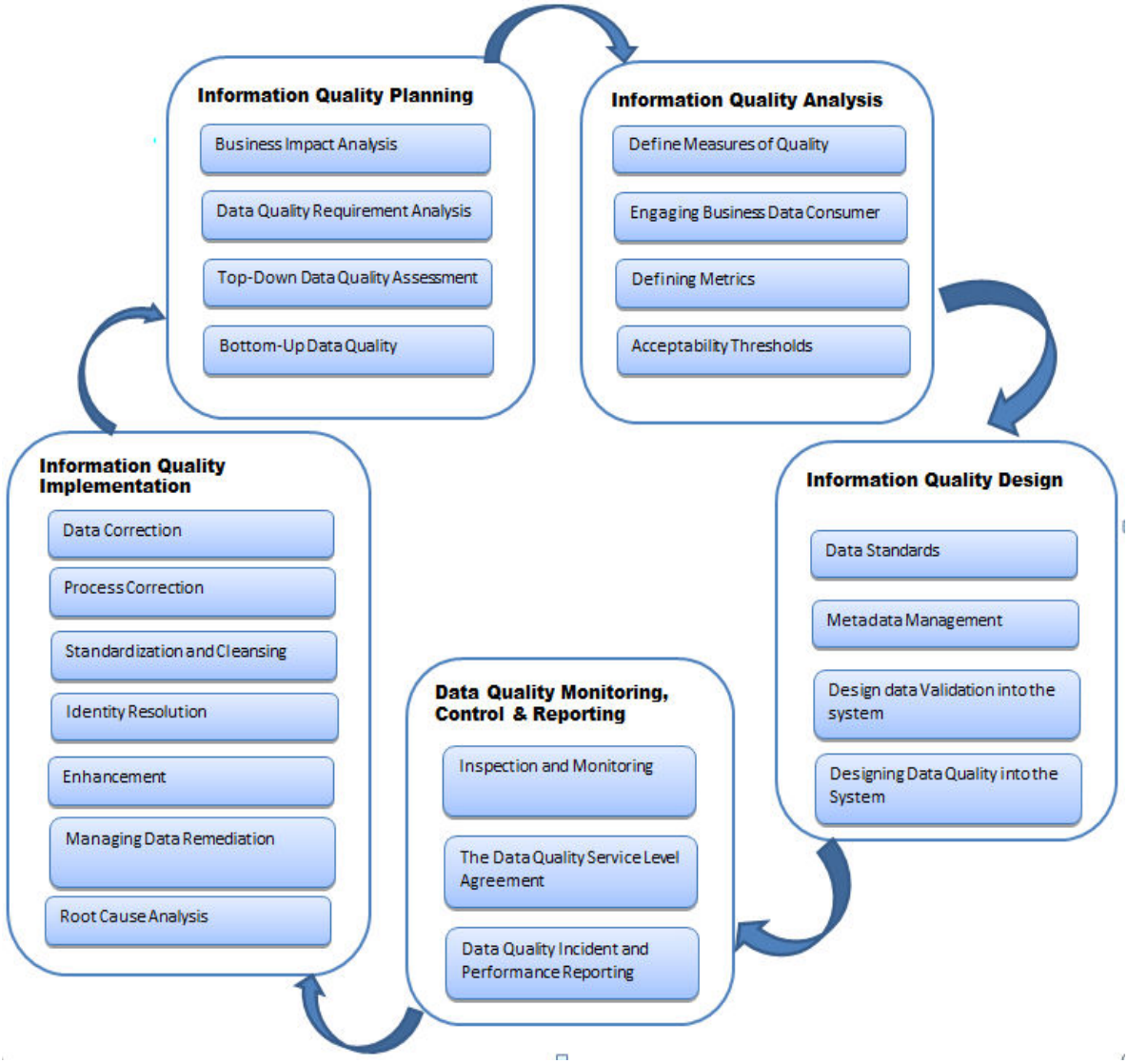

Figure 2.12 Virtuous cycle for information quality, Adapted from Loshin (2011) to group processes or tasks into traditional phases of a system development life-cycle

The table below provides a summarised explanation of each task or process in the Virtuous cycle for information quality:

Table 2.8: Tasks / processes of the virtuous cycle for information quality, adapted from Loshin (2011) to align participants to information stakeholders discussed earlier in this study.

Task / Process	Explanation
Business Impact Analysis	This process enables the information quality custodian to interview information quality consumers and document how selected business processes depend on high information quality. During this process the information quality custodian will note any potential information quality related issues that increase costs, reduce revenues, or introduce inefficiencies or delays in business activities.
Data Quality Requirement Analysis	During this process, the data quality custodian will summarize information quality expectations for consumed information based on the business impact analysis. Information sets will be identified and targeted for assessment, and specific dimensions of information quality will be isolated for review. The selected dimensions will be used to list specific business measures (e.g. Sales, stock) that will be evaluated in relation to the business impacts.
Bottom-Up Information quality assessment	This is a bottom-up, empirical approach to identifying information quality related issues. The information custodian makes use of data profiling and other statistical analysis techniques to identify potential anomalies which are documented in preparation for review with information consumers.
Top-Down information quality assessment	All anomalies discovered in the previous process are reviewed with information consumers. During this process, information custodians work with information consumers to differentiate between relevant and irrelevant issues, prioritize issues based on business impact, and explore different strategies for remediation.
Defining Measures of quality (Business rules)	Business rules are defined to support the information quality management process. The same business rules used in establishing the initial levels of quality can also be used on an on-going basis to inspect information quality at different stages of the information value chain.
Engaging the information consumer	Once the business rules have been defined, information consumers can be brought back in the process to validate them in relation to business impacts.
Defining Metrics	Having identified the information quality dimensions that are relevant to the business processes and identified the specific business rules, the information custodian will then define specific reportable metrics that can be presented information consumers and information managers.
Acceptability Thresholds	In this process, the information custodian request acceptability thresholds from information consumers. The acceptability threshold is the point at which noncompliance with the information consumer's expectations may lead to potential business impact.
Information Standards	This process describes policies and procedures for defining rules and reaching agreement about standard data elements. It is most needed when different business applications share the same information.
Metadata management	This refers to an on-going process which aims to establish an enterprise business metadata management to facilitate the desired level of standards across the organisation. Metadata documentation will include: business uses of information, information definitions, information quality characteristics, inspection and monitoring locations within the information value chain, etc.

Task / Process	Explanation
Data Validity rules	This process aims to transition away from a reactive approach that remediates information quality issues once they have been highlighted by the information consumer. It seeks to engineer controls into the application development process so that information errors can be identified and addresses as they occur.
Inspection and Monitoring	This process involves defining the information quality inspection routines, which may include both automated and manual processes. Inspection procedures are defined for each relevant information quality dimension and customized for each system as appropriate. The availability of business rules for validating information against defined expectations is the basis for information quality inspection and monitoring.
Information Quality Service Level Agreement	This process specifies information consumer's expectations regarding measurable aspects of one or more information quality dimensions.
Information quality incident and performance reporting	This is the process of reporting and tracking information quality issues and corresponding activities.
Designing Information quality into the system	This process incorporates information validation and information inspection and reporting into business processes and corresponding business applications, by adjusting the organisation's system development life cycle (SDLC) to include solicitation of data requirements.
Managing information remediation	This is the mechanism for managing the tasks performed to remedy any critical issues.
Root cause analysis	In this process, the information custodian seeks to identify the source of the introduction of the error, as opposed to just correcting it.
Data Correction	This process corrects information issues identified.
Process correction	This process seeks to determine how processes can be improved to reduce the introduction of errors.
Standardisation and cleansing	This process incorporates working with the data standards and metadata staff to define rules and use tools for standardising information.
Identity resolution	This process seeks to uniquely identify key data concepts within and across different systems
Enhancement	This process focuses on enrichment of information.

2.13. Information quality management in the age of big data

Big data is a large amount of structured, semi-structured, unstructured and raw data in many different formats (Friedman, 2014). It's a paradigm shift in how organisations manage data assets. This means a paradigm shift in how and where data is collected from, how data is analysed and how organisations monetise the insights from the analysis (Kimball et al., 2013). Organisations need to review their data architectures to include several capabilities

that have gained momentum as big data continues to grow (Mohanty et al., 2013). Big data is not the focus of this study and therefore big data capabilities and use cases will not be looked at in detail. However, this section was included in the study as big data has an impact on how organisations manage information quality and how they will consider this in future.

With the advent of big data, organisations produce far more data than they did in the past. Friedman (2014) believes that, as more data is produced by organisations, the amount of bad data also increases. This is especially the case with unstructured data such as social media data which is considered a big data use case. Much of big data is of unknown provenance, which is outside the control of organisations.

From an information value chain perspective, big data is a paradigm shift for all primary and support activities of the information value chain. Information is flowing very fast through the enterprise and it is harder to control. However, the aim of the information value chain, which is to increase the usefulness of information for the information consumers, remains the same. Therefore, proven principles of information quality management should still be valid in the age of big data. In addition to the principles mentioned in the information management section, Friedman (2014) suggests that organisations optimise their business processes, remain compliant to regulations, improve their information governance competency and trust their information in order to cope with information quality in the age of big data.

2.14. Theoretical conceptual framework

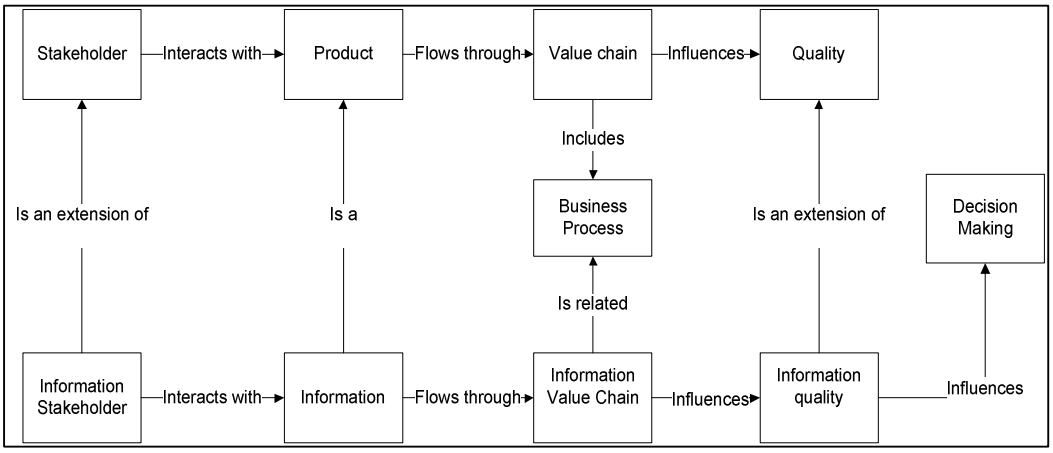

Figure 2.14 Theoretical conceptual framework based on the literature

The framework provided in above image is derived from the literature and it highlights the relationships between some important concepts in this study.

This conceptual framework is based on the concept that stakeholders interact with products of a manufacturing chain. These products flow through a value chain while being manufactured, and the later influences the quality of the product.

This concept is then applied to information systems where an information stakeholder interacts with information products, which flow through and information value chain. The later influences the quality of the information product, which in turn influences decision making.

The following figure is the theoretical conceptual model that will be used in this study. It represents the unit of analysis and unit of observation that will be discussed in the next chapter. It highlights interaction between the information stakeholder and information, which flows in and out of business processes. The framework questions the quality of information that the information stakeholder interacts with, as well as how that affects his decision making.

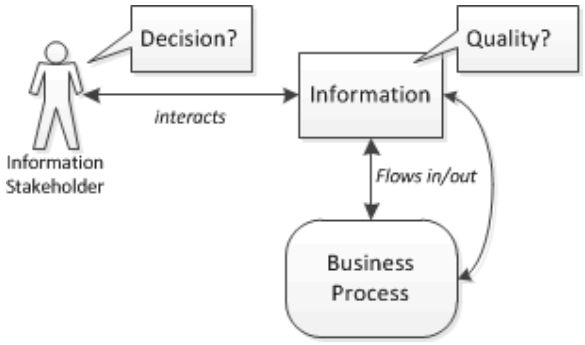

Figure 2.14 The information quality framework proposed for this study

2.15. Conclusion

In this chapter, the literature on information quality was reviewed and discussed. Foundational terms for this study such as data, information and the value chain were introduced and discussed. The concept of value chain was then applied to information management thereby introducing the concept of information value chain. Different information quality dimensions, frameworks and stakeholders were discussed. The different perspectives of information as a product and information as a service were also discussed, before a discussion on the measurement and the management of information quality. The chapter was concluded with 2 conceptual frameworks, one that highlights the relationships between the important concepts of this study drawn from the literature and other representing the units of analysis and unit of observation for the study which will be discussed further in chapter 3

CHAPTER THREE - RESEARCH METHODOLOGY

3.1. Introduction

This chapter highlights the methodological choices followed in this study. It is structured based on the layers of research design framework, also known as the research onion suggested by Saunders and Tosey (2012). The research onion is presented on the next page:

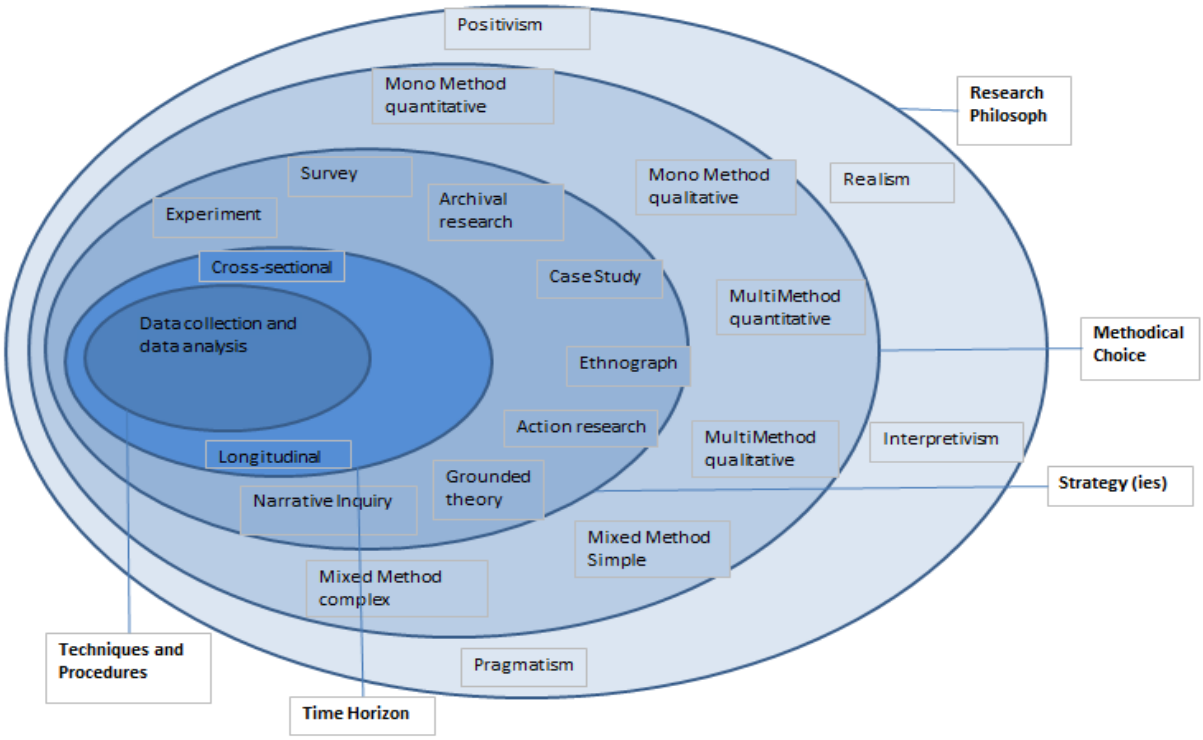

Figure 3.1 The research Onion, adopted from Saunders and Tosey (2012)

3.2. Research philosophy

This is an interpretive study that seeks to understand and explain the subjective topic of information quality within a chosen retail organisation. Saunders and Tosey (2012) believe that a researcher reflects the philosophy of interpretivism when he or she is more concerned with gathering insights into subjective meanings than providing law-like generalisations. There are other research philosophies available such as positivism where the researcher is concerned with observing and predicting outcomes; realism which says that reality exists independent of the mind; or pragmatism where researchers believe that the importance of research is in the practical consequences of the findings. However, for this study, based on the researcher's assumptions about knowledge and how it is obtained and since the researcher is more concerned with gathering rich insights into a subjective topic, interpretivism is the most suitable philosophy. For research to be significant, it has to be

trustworthy (De La Harpe, 2008). Carr and Kemmis (1986) suggest that an interpretive research account needs to be coherent and must pass the test of participant confirmation. Participant confirmation is an important aspect of this research. The focus of this study is on conducting research amongst people rather than objects as the researcher would like to adopt an empathetic stance so as to understand the respondents' social world and the meaning that they attach to it from their perspective

3.3. Methodical choice

The study makes use of a qualitative research paradigm. Niemann (2005) describes qualitative research as any type of research that produces findings not arrived at by statistical procedures or other means of quantification. Data was collected through in depth interviews. Boyce and Neale (2006) define an in depth interview as a qualitative research technique that involves conducting intensive individual interviews with a small number of respondents to explore their perspective on a particular idea. In this study, detailed information about participants' thoughts on information quality for the chosen business process will be collected. In depth interviews were chosen as a method for collecting data in this study because they provide much more detailed information than what is available through other data collection methods (Boyce & Neale., 2006).

3.4. Research strategies

A case study research strategy was adopted for this study. A case study is defined as an empirical method aimed at investigating contemporary phenomena in their context (Runeson & Host., 2008). There are other research strategies related to case studies. For example, surveys, which Runeson and Host (2008) define as the collection of standardised information from a specific population, usually by means of a questionnaire or interview. Experiment in another example and it is characterised by measuring the effects of manipulating one variable on another variable. Action research is another example of a research strategy within which the researcher works with practitioners to bring about organisational change (Boyce & Neale., 2006). A case study was chosen as the preferred strategy because the main research question being posed is a "How" question. Yin (2003) supports this argument when he states that a case study has a distinct advantage in a situation when a "how" or "why" question is being asked about a set of events over which the researcher has little or no control. Case studies have been used in the past in information systems research in general (Runeson & Host., 2008), and information quality in particular (De La Harpe, 2008). Another reason for choosing the case study is that they are generally associated with interpretive research whenever a qualitative research paradigm is adopted with a subjective stance. Although Boyce and Neale (2006) warn that case studies can also be used in positivistic research. Although the researcher's role is that of an observer, through case study research

he becomes part of the real-world. The case study chosen was the ordering business process, also known as end to end supply chain process of a major retail organisation. Due to the size and complexity of the ordering process and the detailed nature of the retail environment, only the import part or the process was considered.

3.5. Time horizon

This research was conducted from 2012 to 2014, with all interviews taking place in 2013. It was undertaken to answer information quality questions at that particular time, which makes it cross-sectional research as opposed to a longitudinal study which necessitates the collection of data for an extended period of time.

3.6. Data collection

The researcher conducted in depth interviews with different data stakeholders that are involved in the ordering business process. Although the researcher has formed some pre-conceived ideas about data quality from analysing the literature and from being part of the chosen organisation used as a case study, he assumed the role of a complete observer in this study. De La Harpe (2008) highlights three other types of observer roles in observational research, namely: Complete participant, participant as observer and observer as participant. In this study, the researcher seeks to understand the issues experienced by other participants without actually interacting with the data objects. This interaction with data stakeholders was through in depth interviews on the quality of information in the chosen case study. De La Harpe (2008) states that interviews are good methods to find out what information stakeholders think, believe and why they react in a certain way. Data was also collected from organisational documents such as business process documentation and departmental training manuals. The next section discusses the collected data that will be analysed.

3.7. Data analysis

In this study, the researcher used two data quality frameworks as a theoretical lens to guide the empirical data analysis. The choice of these frameworks was based on the preamble that information quality is a multidimensional concept. The categorical information quality framework by Strong et al., (1997) which insists on the importance of information quality categories and dimensions, as well as the Semiotics information quality framework by Shanks and Corbitt (1999) which is based on context from semiotic theory, which uses symbols to transfer knowledge were the two chosen frameworks. These frameworks and other frameworks available in the literature have been discussed in the previous chapter.

Finally the data was analysed according to the theoretical conceptual framework based on the literature reviewed. This framework was discussed at the end of the previous chapter. The unit of analysis and unit of observation for this study are discussed in the next section.

3.8. Unit of analysis and unit of observation

This study seeks to understand the information quality perceptions of stakeholders as they interact with information throughout the chosen business case. Their perceptions may be influenced by factors related to their environment as well as organisational structures. As defined in the previous chapter, information stakeholders are those with interest in information with which they interact in various roles. These roles are: information producers, information custodians, information consumers and information managers. Information is produced and consumed by business processes and these stakeholders interact with information throughout the organisation's information value chain.

In this study, the environment in which information stakeholders operate is also of importance and it is believed that the way in which they operate within that environment cannot be isolated from influences of environmental factors. The unit of observation is the information stakeholders' perceptions of information quality at the different touch points of the business processes. The unit of analysis is the information quality according to the dimensions proposed by the selected frameworks. The research design for this study is proposed in the next section.

3.9. Research design

The research philosophy adopted in this study was interpretive. The purpose of the research is to understand information stakeholder's perceptions of information quality as they interact with information through the information value chain. Two information quality frameworks were used to analyse the data. The interpreted results should improve the understanding of how information quality is perceived by different stakeholders involved in the same business process. In turn, this improved understanding could contribute towards an improvement of the quality of information for the organisation chosen as a case study. The case study is a business case describing the end to end supply chain process of a retail organisation. The level of analysis was information stakeholders in the context of the chosen organisation. The researcher collected data through in depth interviews with identified information stakeholders.

The research process follows the steps below:

- Identify an area of study and describe the purpose of the study

- Conceptualise the problem by describing the background
- Describe the philosophical perspective adopted for the research
- Analyse the relevant literature to establish the current state of the research
- Conduct the research by selecting a site, collecting data and transcribing interviews
- Analyse and interpret the collected data to make sense of it
- Formulate the findings, draw conclusions and highlight significances and limitations of the study.

3.10. Ethics

This study followed ethical guidelines, as specified by the CPUT code of practice on ethical standards. Permission to conduct research was obtained from the organization chosen for the case study. Informed consent was obtained from all participants in the study. The business case is presented and analysed in the next chapter.

CHAPTER FOUR - BUSINESS CASE

4.1. Introduction

In this chapter, the business case used for the empirical part of the study is discussed. The chosen organization and the business process used for the business case are introduced. The methods used for data collection, the researcher's role in the study, the selection of the research site and the period of the study are discussed before the chosen information quality frameworks are used to present the data collected. The collected data is then analysed, before findings are summarised.

4.2. Description of the business case

In the next sub section the organisation used as the business case is described.

4.2.1. The retail organisation

As mentioned above, the chosen organisation for this study is a retailer. This business offers a unique blend of merchandise which includes fashion, foods, beauty products and homeware products to millions of customers in several countries. This business always puts its customers first and is very passionate about quality, service, innovation and sustainability. Its support services, which include information technology services, are focused on partnering with the merchandise areas of the business to ensure that the company fulfils its promises every day.

The ordering business process also known as end to end supply chain process of this major retail organisation was chosen as the case study. Due to the big size and complexity of the process as well as the limited amount of time available to conduct this study, it was decided to only focus on the import component. This process involves several business information stakeholders from different departments in the organization, namely; the buying department, sourcing department, merchandise planning department, international supply chain department and the finance department. The process also involves stakeholders that are external to the organisation, such as suppliers and logistic companies.

To support this process, a number of information systems are used. These include an import trade management system, a merchandising system, logistic systems, airfreight tracking systems, a financial information system, a warehousing or distribution centre management system, planning systems, business intelligence, and excel spreadsheets. The following sections describe the different areas of the business involved in the business case.

4.2.2. Buying department

Buying is responsible for the procurement of products as per the organisation's strategy. They work with a design team to understand seasonal trends, customer preferences and develop them into relevant products for the company. They are responsible for building the required products ranges, negotiating with suppliers and dealing with the administration required for this, while ensuring that the ranges meet sales and profit budgets.

4.2.3. Sourcing department

Sourcing is responsible for identifying appropriate factories and suppliers, while supporting the organisation's product strategy. They manage the chosen business processes critical path to ensure on-time delivery, and build strong relationships with suppliers and key stakeholders. They also perform all the administration required in the imports business process.

4.2.4. Merchandise planning department

Planning works with buying to build a merchandise range that delivers the business's key performance indicators (KPI). They are responsible for delivering profitable growth through effective financial and product planning as well as management of stock. They perform planning tasks such as production planning, allocations as well as the administration required in these tasks.

4.2.5. International supply chain department

This department is part of the broader supply chain which deals with moving stock from suppliers to final delivery destinations. In the chosen business process, the international supply chain department works closely with the revenue services and deals with in-house customs clearance of all goods imported by the buying department.

4.2.6. Finance department

Like in any organisation, the finance department is responsible for managing financial resources. For the retail organisation chosen in this study, managing financial resources includes the following tasks:

- Performing planning and management accounting
- Performing revenue accounting
- Managing taxes
- Managing and consolidating international funds
- Managing treasury operations
- Managing internal controls

- Performing general accounting and reporting
- Processing the payroll
- Processing accounts payable and expense reimbursements
- Managing stock integrity
- Managing property
- Managing financial information

Within the chosen business process, the finance department is in charge of ensuring that the suppliers are paid.

4.2.7. Suppliers

Only international suppliers, meaning suppliers that are located outside the country where the chosen organisation is based, of fashion products are involved in the chosen business process.

4.2.8. Logistic company

These are external companies that this organisation works with to facilitate transportation of goods from the supplier's country of origin to where the organisation is based.

4.3. Data collection and analysis

Different business stakeholders were interviewed to collect data on the chosen business case. The research questions were taken into consideration at this stage and a summary of the main findings is presented at the conclusion of the chapter. To collect the data during the interviews, the researcher used a very collaborative and visual approach. Sticky notes and A1 paper sheets were used to plot all issues as well as the information quality dimension applicable to the issue. The image below shows a picture of what the collected data initially looked like.

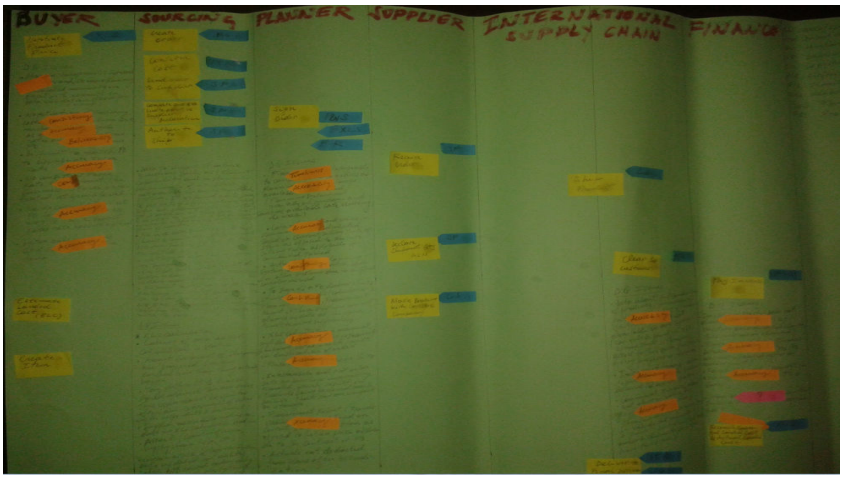

Figure 4.3 A picture of the organising of the collected data

The approach used to collect and analyse the data considered the pre-amble that information quality is complex and that a multi-dimensional view will provide better insight. Two frameworks were used to present and analyse the data collected, namely; the categorical information framework by Strong et al., (1997) and the Semiotics information quality framework by Corbitt (1999). Both frameworks were found suitable for the study as they support the multi-dimensional view of information quality. Data for each information quality dimension discussed in chapter 2 is presented and categorised in the form of a matrix or summary Table. Each row present a summary of responses collected from business information stakeholders for a specific information quality dimension. The columns on the other hand present each stakeholder responses across all information quality dimensions.
A summary of the findings is presented at the conclusion of the chapter. Similarities and differences between the different perspectives are considered to establish the usefulness of the multi-dimensional view of information quality.

4.4. The role of the researcher

Permission for the study was obtained from the managers of all departments involved who explained to the other employees the potential benefit that the organisation can gain from the research findings. Although the perceptions of the respondents were sought, the focus was always on the information and how that supports or prevents the respondents' actions.

The researcher and the respondents who participated in this study work for the same organisation. Respondents were from different departments within the same retail business unit and the researcher works in the information technology department. Being in that position, the researcher had the advantage of talking to people with whom he had built a relationship in the past and, therefore, the respondents trusted him. In all the cases he explained to the participants the purpose of the study and that the results will not compromise their work positions. He also explained to them that their responses will be

coded and therefore be anonymous. The researcher was perceived by most respondents as someone who would contribute to the resolution of some of the information quality related issues that they experience. As a result, respondents were willing to share a lot of information related to information quality frustrations that they experience. Where there was a possibility of misunderstanding, the researcher asked his respondents specific questions to explain the reason for their actions. In most cases, the researcher interviewed more than one respondent from the same department to ensure that the study does not represent an individual's perspective on information quality.

4.5. Selection of the research site

The research site was selected as a good example of a modern, retail, information driven organisation in which business processes generate large amount of data, and should, therefore, face sufficient amounts of information quality challenges. As highlighted in the previous section, the researcher works for the organisation, therefore, access to the organisation was not a major challenge for him once he obtained permission to conduct the empirical part of the research study.

4.6. Period of study

The period of study was from 2012 to 2014, with the majority of interviews taking place during 2013. This was not a longitudinal study but rather a cross-sectional study conducted during that period. Due to the familiarity of the researcher with the organisation used as a case study, some of the interviews were scheduled on an ad-hoc basis whenever the research felt he needed more information.

4.7. Mapping of the business process

To get an understanding of the business process as well as the information quality related issues that it experiences, the researcher interviewed representatives from each of the information stakeholder groups involved. From the data collected, the author mapped his understanding of the business process on the following page.

As highlighted above, this study only focuses on the import section of the ordering process. There is a different process in place for orders sourced from local suppliers; however, that was excluded from the scope of this research.

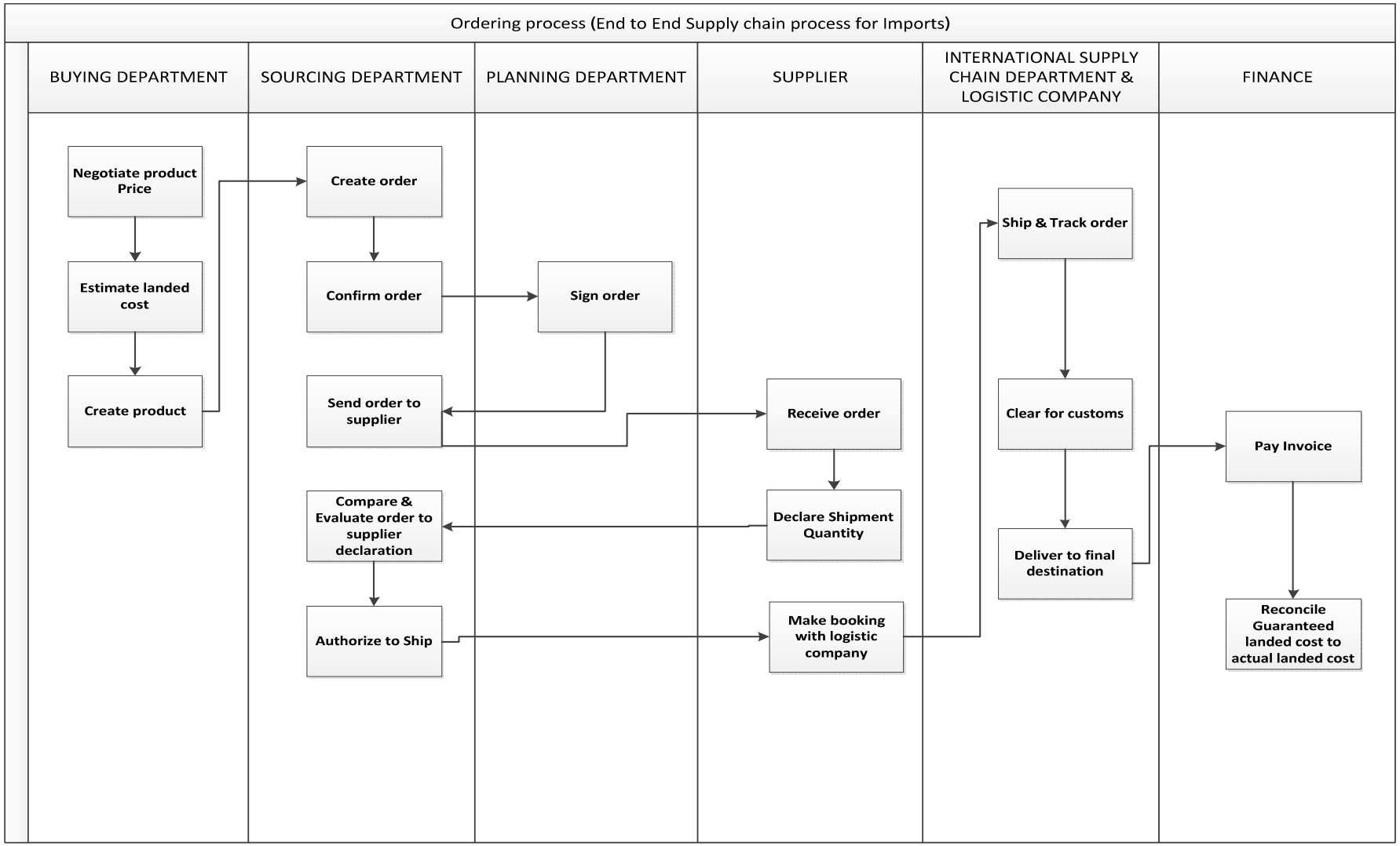
Ordering process (End to End Supply chain process for Imports)
BUYING DEPARTMENT
SOURCING DEPARTMENT
PLANNING DEPARTMENT
SUPPLIER
INTERNATIONAL SUPPLY CHAIN DEPARTMENT & LOGISTIC COMPANY
FINANCE
Negotiate product Price
Estimate landed cost
Create product
Create order
Confirm order
Send order to supplier
Compare & Evaluate order to supplier declaration
Authorize to Ship
Sign order
Receive order
Declare Shipment Quantity
Make booking with logistic company
Ship & Track order
Clear for customs
Deliver to final destination
Pay Invoice
Reconcile Guaranteed landed cost to actual landed cost

Figure 4.7 Ordering business process for imports

The following section is a short description of the business processes depicted in above Figure.

4.8. Description of the business process

The buying department kicks off this business process by negotiating a product price with the supplier. Once the price has been agreed upon, the buyer estimates a landed cost using an import management system. The last involvement from the buyer in this process is to create the product / item in the merchandising system.

Once the item has been created in the merchandising system, the sourcing department can create a purchase order in the same system. They will then confirm the landed cost in the import management system, before sending the purchase order to the planner who will sign it off. Once signed off, the planner sends that purchase order to the supplier in an electronic form through a supplier portal interface.

As soon as the supplier receives the order, he will declare the number of items that he is able to ship and send that information back to the retail organisation through the same interface used to send the purchase order. The sourcing department will then compare the quantity ordered versus the quantity declared by the supplier before authorizing the shipment, again via the supplier portal.

After receiving authorisation to ship the products, the supplier books a shipment date with the logistic company. The latter will then ship the products, with the shipment recorded in the import management system. Once the goods have been shipped and received in the country of the retailer, the international supply chain department is involved with the clearance of goods for customs. Once cleared, the goods will only then be delivered to their final destination with the delivery recorded in a warehouse management system. Depending on the payment terms agreed with the supplier, the finance department will pay the invoice and reconcile costs.

The next table highlights different points in the business process where stakeholders interact with the process and the main information products that they produce, manage or consume.

Table 2.1: Business processes and points of interaction

Information Product	Information Stakeholder	Role	Source Process	Destination Process	Example of information quality issue
Product price	Buying department	Information producer	Negotiate product price	Create order	Product price captured incorrectly
Order information	Sourcing department	Information producer	Create order	Authorize to ship	Incorrect port loaded for a purchase order
	Planning department	Information manager			
	Supplier	Information consumer			
Shipment information	Sourcing department	Information producer	Authorize to ship	Deliver to final destination	Shipment date incorrectly estimated
	Supplier	Information producer			
	International supply chain	Information manager			

4.9. Findings from the business process understanding

Most activities in this business process are managed in an automated way through sustainable information systems. However, a number of activities still rely on excel spreadsheets.

Every business stakeholder role involved in this business process experiences information quality related issues. Whenever data quality issues are experienced, the business process is directly affected. For example, if supplier information is not up to date, the finance department might not be able to pay the supplier.

Not all stakeholders agree on the causes of information quality issues experienced. Some believe that these issues are caused by business stakeholders not following the process and others believe that information quality issues are a result of bugs in information systems.

The next section organises and presents these issues using two of the information quality frameworks highlighted in chapter 2: Categorical information quality framework and the semiotic information quality framework.

4.10. Categorical information quality matrix

This section analyses the data collected using the categorical information quality matrix. This matrix is based on Table 2.1 in chapter 2, where the importance of information quality categories and dimensions is emphasised.

4.10.1. Intrinsic information quality category

Accuracy was the dimension found with the highest number of information quality challenges. Seventeen (17) distinct issues where highlighted by business information stakeholders for the accuracy dimension. No issues were highlighted for the objectivity and reputation dimension, while one issue was highlighted for believability. Manual processing was often highlighted as the cause of information accuracy issues. All groups of business information stakeholders that were interviewed by the researcher raised at least 2 information accuracy related issues that they experienced in of the chosen business process. Some issues were shared by different groups of information stakeholders.

Listed below are some examples of issues by dimension for the intrinsic information quality category:

- Accuracy: The most common issues were related to manual calculations built into excel spreadsheets with no data validations, therefore, introducing room for error. These were followed by a number of concerns related to incorrectly captured data such as incorrect port on the purchase order, incorrect shipment date, incorrect tariff applied, and so forth
- Objectivity: There were no issues or concerns related to objectivity
- Believability: Whenever business stakeholders are faced with multiple foreign exchange rates for the same purchase order, they are not sure which exchange rate is to be used
- Reputation: There were no issues or concerns related to reputation.

Listed below are some examples of issues by stakeholders for the intrinsic information quality category:

- Buyers: Respondents from this group highlighted issues caused by the use of excel spreadsheets to manually calculate some important business measures as well as the multiple foreign exchange rates available for the same purchase order
- Sourcing: Respondents from this group also raised issues with the use of excel spreadsheets with no validations, incorrectly captured data and data that was not up to date
- Planners: Respondents from this group raised issues about incorrectly calculated business measures due to non-existence of data needed for the correct calculation, as well as incorrectly captured data

- International supply chain: respondents from this group raised issues about incorrect tariff applied for custom clearance due to the use of paper based data sources and manual process followed to access information
- Finance: Respondents from this group raised issues about data not being up to date, incorrect foreign exchange rates applied to purchase orders and mismatches in values captured on purchase orders versus invoices.

The main finding with the intrinsic data quality dimension is that errors occur due to the use of manually captured data in excel spreadsheet in addition to data in automated systems.

4.10.2. Accessibility information quality category

No issues were found with access security, while 5 issues were highlighted with the accessibility dimension. Lack of computerised information to be used for decision making was the most cited in this category, followed by a lack of an integrated view of information. With the exception of finance, every group of business information stakeholder interviewed for the study highlighted an issue related to accessibility of information. This seemed to be a bigger issue for stakeholders operating at a strategic level of the organisation as they felt that it takes too long to access information needed for a strategic decision.

Listed below are some examples of issues by dimension for the accessibility information quality category:

- Accessibility: This was the only dimension in this category for which issues were raised. The issues were mainly related to the lack of or limited access to integrated information. As a result, business stakeholders spend a lot of time trying to get to the information needed for decision making. There were also a couple of issues where the information was not available systematically
- Access security: There were no issues or concerns related to access security.

Listed below are some examples of issues by stakeholders for the accessibility information quality category:

- Buyers: Respondents from this group raised issues about the amount of time it takes them to arrive to the information that they actually need for decision making
- Sourcing: Respondents from this group raised issues about the lack of information in an integrated system to answer their strategic business questions. They also highlighted an example of supplier data that is not available in most systems

- Planners: Respondents from this group highlighted issues about systems not being available during weekends and certain times of the day
- International supply chain: Respondents from this group raised issues about the lack of computerised information, and the excessive use of excel spreadsheets
- Finance: Respondents from this group did not raise issues in this category.

The main finding with regards to the accessibility quality dimension is the lack of an integrated view of data across the different systems.

4.10.3. Contextual information quality category

In this category, no concerns were found with relevancy, value-added, and completeness dimensions. However, 4 issues were highlighted regarding the timeliness dimension and 2 distinct issues were found with the 'amount of information' dimension. With the timeliness dimension, the frequency of information updates was the most cited issue across business stakeholders. As for the amount of information, it was either too much information to go through before arriving to a decision, or simply too little or an insufficient amount of information available in the information system. Issues related to the amount of information dimension were very similar to those in the accessibility dimensions.

Listed below are some examples of issues by dimension for the contextual information quality category:

- Relevancy: There were no issues or concerns related to relevancy of information
- Value-added: There were no issues or concerns related to this dimension
- Timeliness: Issues related to this dimensions where about the frequency at which data is updated. Lack of real-time data for supplier margins, Insufficient or lack of real-time data updates in logistic system, as well as unavailability of planning information when needed for decision making, are some of the examples
- Completeness: There were no issues or concerns related to this dimension
- Amount of information: The large amount of time it takes to arrive of information needed for decision making as well as the lack of sufficient information to support strategic decisions were the main issues raised for this dimension.

Listed below are some examples of issues by stakeholders for the contextual information quality category:

- Buyers: Respondents from this group raised issues about large amount of time it takes to arrive of information needed for decision making. They referred to this as the lack for exception reporting
- Sourcing: Respondents from this group raised issues about the lack of real-time data updates, the large amount of time it takes to arrive of information needed for decision making and the lack of information needed to support strategic decisions
- Planners: Respondents from this group raised issues about the unavailability of planning data and planning systems when needed
- International supply chain: Respondents from this group did not raise issues in this category
- Finance: Respondents from this group did not raise issues in this category.

The main finding for the contextual quality dimension is that the amount of data on which a decision is based is either too much or too little.

4.10.4. Representational information quality category

A concise representation of the information was not highlighted as an issue. However, one issue was highlighted for interpretability, another one on ease of understanding, and consistent representation was the biggest issue in this category with 4 distinct issues. Planners, the international supply chain and finance department all had issues with the consistent representation of information. These often had to do with different systems being out of synchronisation.

Listed below are some examples of issues by dimension for the representational information quality category:

- Interpretability: The difficulty to tell whether there has been profit or loss on a particular purchase order due to multiple exchange rates, was the issue highlighted for this dimension
- Ease of understanding: The difficulty to match purchase orders to the applicable foreign exchange rate was the only issue raised for this dimension
- Concise representation: There were no issues or concerns related to this dimension
- Consistent representation: Issues related to this dimension were about data in planning systems not matching data in other systems, such as the data warehouse and the merchandising system, due to timing differences and different calculations used. There were also issues raised about values on invoices not matching their related purchase orders.

Listed below are some examples of issues by stakeholders for the representational information quality category:

- Buyers: Respondents from this group raised issues about the difficulty to match purchase orders to the relevant foreign exchange rates
- Sourcing: Respondents from this group raised issues about the lack of clear visibility of profit or loss and the difficulty to match purchase orders to the relevant foreign exchange rates
- Planners: Respondents from this group raised issues about data in planning systems not matching other systems due to timing issues or different calculations applied to derive same business measures
- International supply chain: Respondents from this group raised issues about values reflected on purchase orders not matching their related invoices
- Finance: Respondents from this group raised issues about values reflected on purchase orders not matching their related invoices.

The main finding relating to the representational quality dimension is that data are not represented in the same manner across systems.

The table below shows detail responses from information stakeholders captured in the categorical information quality matrix.

Table 4.2: Responses from information stakeholders captured in the categorical Information Quality Matrix

		Business Information Stakeholders				
DQ Category		**Buyer**	**Sourcing**	**Planner**	**Int. Supply Chain**	**Finance**
Intrinsic Information Quality	**Accuracy**	Forward Margin calculated manually in Excel with No built in data validation. That introduces room for error. Multiple Forex rates available for same purchase order	Some information needed for strategic questions is manually controlled in excel spreadsheets with no validations Critical path management of order / order tracking done manually in Excel 60 % of shipments estimated for an earlier time than they actually happen Supplier master data not up-to-date in Excel calculator and planning system Incorrect port loaded for a purchase order, therefore sourcing cannot proceed with order	Commitment rand values in planning systems & data warehouse procurement report Rand values, calculated based on current selling price as opposed to the future selling price applicable for the committed period Incorrect Shipment dates generated from XLS calculator Duplicates Purchase orders created Incorrect Price and Quantity captured on Purchase order Incorrect In-store data and Not-before date generated from XLS calculator	Incorrect tariff applied as a result of manual process to access the information Commercial invoice does not match purchase order (issue occurs for at least one or 2 shipments per week)	Supplier Master data not up-to-date Price & Quantity differences between Invoice and Purchase order Incorrect Forex applied to a purchase order
	Objectivity					
	Believability	Multiple Forex rates available for same purchase order				

		Business Information Stakeholders				
DQ Category		**Buyer**	**Sourcing**	**Planner**	**Int. Supply Chain**	**Finance**
	Reputation					
Accessibility / Interactional Information Quality	**Accessibility**	It takes too long to get to information needed for a decision. No exception reporting	Information needed to answer strategic questions not found in one integrated system Local currency of the Supplier's country of origin not available in most systems	Planning systems not accessible on Sundays and sometimes accessible late during the week	Information used for tariffing (e.g: full composition of products, percentage of fabric, country of manufacture, etc) not available in any automated system. Can only be found on XLS spreadsheet	
	Access security					
Contextual Information Quality	**Relevancy**					
	Value-added					

		Business Information Stakeholders				
DQ Category		**Buyer**	**Sourcing**	**Planner**	**Int. Supply Chain**	**Finance**
	Timeliness		No live data available to compare current vs actual margin for a supplier Excel calculator used for deriving shipment dates is updated manually, not always up to date Some data in Logistic system only updated on nightly batch	Planning data / planning systems not always available when needed		
	Completeness					
	Amount of Information	It takes too long to get to information needed for a decision. No exception reporting	It takes too long to get to information needed for a decision. No exception reporting Not enough information available in systems to support strategy			
Representational Information Quality	**Interpretability**		No clear visibility of profit or loss due to difficulty of applying correct Forex rate			
	Ease of understanding	Difficult to match Purchase order to appropriate Forex rate	Difficult to match Purchase order to appropriate Forex rate			

		Business Information Stakeholders				
DQ Category		**Buyer**	**Sourcing**	**Planner**	**Int. Supply Chain**	**Finance**
	Concise representation					
	Consistent representation			Actual data in Planning systems does not match data in other systems due to different catalogues used To Sign-off a Purchase order, a planner needs to check it against Planning systems, XLS Calculator, and Data warehouse procurement report. Yet, Purchase order data in Planning systems, XLS Calculator, and Data warehouse procurement report not updated at the same time (i.e. timing differences)	Commercial invoice does not match purchase order (issue occurs for at least one or 2 shipments per week)	Price & Quantity differences between Invoice and Purchase order

4.10.5. Summary of findings from a categorical information quality perspective

- Most information quality issues were related to the accuracy dimension, followed by accessibility and consistent representation
- No issues were highlighted regarding the objectivity, reputation, access security, relevancy, value-added, completeness and concise representation dimensions
- The sourcing department, followed by planners and buyers raised the highest numbers of information quality related issues
- International suppliers raised the lowest number of issues but it is also important to note that a lot of data needed to do their job is not available in any automated system.
- Most of the findings relate to the need to use additional, manually captured data in addition to the ones produced by the system
- Data are also not represented in the same manner in the different systems and seems to be not sufficiently integrated
- Data required for decisions is not easily obtainable

4.11. Semiotics information quality matrix

This section analyses the same data collected using the semiotic information quality matrix. This second matrix used to analyse the data collected is based on Table 2.2 in chapter 2.

4.11.1. Syntactic

Not many issues were raised from the syntactic view. Only two of the issues raised had to do with non-conformance to meta data. Those were the incorrectly generated shipment dates and incorrectly calculated commitment rand values for future periods. These issues were raised by the sourcing and planning departments. It seems as if the syntactic quality aspect of information is sufficiently addressed by the systems.

4.11.2. Semantics

Most issues from a semantic view had to do with correctness of the information. Every group of information stakeholder interviewed for this research had examples of information correctness issues that they experienced as part of the chosen business process. Similarly to accuracy in the categorical information quality framework, correctness was the dimension with the highest number of issues in the semiotics information quality framework. Several issues were also highlighted for the non-redundant, unambiguous and complete dimensions. For example, no consistent view of Forex rates across departments in the organisation as each department had its own Forex rate, supplier master data kept in multiple systems and not always updated at the same time.

It follows from this that information stakeholders were not always able to attach sufficient meaning to the data and often had to capture additional data manually.

4.11.3. Pragmatic

Accessibility of information was by far the dimension with the highest number of issues from a pragmatic view. The buying and sourcing department also had issues with some information not being understandable with the most notable challenge being the difficulty of matching a purchase order to its applicable Forex rate, which leads to lack of clear visibility of profit or loss incurred by the department. A couple of issues were also raised mainly by the sourcing department on security, relevance and flexibility of presentation.

Following from the above it is clear that information stakeholders could not always use the data to perform their tasks and had to perform additional data manipulations in order to have useful data.

4.11.4. Social

The buying and sourcing department had examples where the quality of information was not sufficient to support their daily activities. Other examples were provided by the finance and sourcing department where the quality of information was not good enough to support relationships between stakeholders, whether internal or external and to share a good understanding with others.

It follows from the above that information with quality problems influences the relationships between the different departments and results in communication problems.

The following Table below shows detail responses from information stakeholders captured in the semiotics information quality matrix.

Table 4.3: Responses from information stakeholders captured in the semiotics information quality matrix

		Business Information Stakeholders				
		Buyer	**Sourcing**	**Planner**	**Int. Supply Chain**	**Finance**
Syntactic - information conformance to metadata (i.e. database rules)	**Conforming to metadata**		60 % of shipments estimated for an earlier time than they actually happen	Commitment rand values in planning systems & data warehouse procurement report Rand values, calculated based on current selling price as opposed to the future selling price applicable for the committed period		
Semantic - information correspondence to external phenomena	**Complete**		No live data available to compare current vs actual margin for a supplier Excel calculator used for deriving shipment dates is updated manually, not always up to date Some data in Logistic system only updated on nightly batch			
	Unambiguous	Multiple Forex rates available for same purchase order		To Sign-off a Purchase order, a planner needs to check it against Planning systems, XLS Calculator, and Data warehouse procurement report. Yet, Purchase order data in Planning systems, XLS Calculator, and Data warehouse procurement report not updated at the same time (i.e. timing differences)		

		Business Information Stakeholders				
		Buyer	**Sourcing**	**Planner**	**Int. Supply Chain**	**Finance**
	Correct	Forward Margin calculated manually in Excel with No built in data validation (room for error) No way to check that correct Forex rate has been applied	Some information needed for strategic questions is manually controlled in excel spreadsheets with no validations Critical path management of order / order tracking done manually in Excel 60 percent of shipments estimated for an earlier time than they actually happen Supplier master data not up-to-date in Excel calculator and planning system Incorrect port loaded for a purchase order, therefore sourcing cannot proceed with order	Commitment rand values in planning systems & data warehouse procurement report Rand values, calculated based on current selling price as opposed to the future selling price applicable for the committed period Incorrect Shipment dates generated from XLS calculator Duplicates Purchase orders created Incorrect Price and Quantity captured on Purchase order Incorrect In-store date and Not-before date generated from XLS calculator Actual data in Planning systems does not match data in other systems due to different catalogues used	Incorrect tariff applied as a result of manual process to access the information Commercial invoice does not match purchase order (issue occurs for at least one or 2 shipments per week)	Supplier Master data not up-to-date Price & Quantity differences between Invoice and Purchase order Incorrect Forex applied to a purchase order

		Business Information Stakeholders				
		Buyer	**Sourcing**	**Planner**	**Int. Supply Chain**	**Finance**
	non-redundant	No Consistent view of Forex Rates throughout the business - Forex rates negotiated at group level	Supplier master data not up-to-date in Excel calculator and planning system Information needed to answer strategic questions not found in one integrated system.	Actual data in Planning systems does not match data in other systems due to different catalogues used To Sign-off a Purchase order, a planner needs to check it against Planning systems, XLS Calculator, and Data warehouse procurement report. Yet, Purchase order data in Planning systems, XLS Calculator, and Data warehouse procurement report not updated at the same time (i.e. timing differences)		
	Meaningful					
Pragmatic - information worth (importance) for use	**Accessible**	It takes too long to get to information needed for a decision. No exception reporting.	Local currency of the Supplier's country of origin not available in most systems Information needed to answer strategic questions not found in one integrated system. No live data available to compare current vs actual margin for a supplier	Planning systems not accessible on Sundays and sometimes accessible late during the week	Information used for tariffing (e.g: full composition of products, percentage of fabric, country of manufacture, etc) not available in any automated system. Can only be found on XLS spreadsheet.	
	Presentation (flexibly presented)	It takes too long to get to information needed for a decision. No				

		Business Information Stakeholders				
		Buyer	**Sourcing**	**Planner**	**Int. Supply Chain**	**Finance**
		exception reporting.				
	Understandable	Difficult to match Purchase order to appropriate Forex rate	No clear visibility of profit or loss due to difficulty of applying correct Forex rate Difficult to match Purchase order to appropriate Forex rate			
	Secure		Some information needed for strategic questions is manually controlled in excel spreadsheets with no validations			
	Relevant		Not enough information available in systems to support strategy			
	Valuable					

		Business Information Stakeholders				
		Buyer	**Sourcing**	**Planner**	**Int. Supply Chain**	**Finance**
Social - Represents Beliefs, expectations, commitments, contracts, law, culture …	**Quality of information sufficient to support daily activities**	Difficult to match Purchase order to appropriate Forex rate	Local currency of the Supplier's country of origin not available in most systems Not enough information available in systems to support strategy			
	Quality of information sufficient to support communication with others					
	Quality of information sufficient to support relationships between stakeholders		Supplier master data not up-to-date in Excel calculator and planning system Local currency of the Supplier's country of origin not available in most systems No live data available to compare current vs actual margin for a supplier			Supplier Master data not up-to-date
	Possible to interpret information to support needs;					

		Business Information Stakeholders				
		Buyer	**Sourcing**	**Planner**	**Int. Supply Chain**	**Finance**
	Extent it is influenced by organisational beliefs, culture and expectations;					
	Ability to share quality understandings with others		No clear visibility of profit or loss due to difficulty of applying correct Forex rate			

4.11.5. Summary of findings from a semiotics information quality perspective

- The 'Correct' dimension was associated with the highest number of information quality issues, followed by 'Accessible' and 'non-redundant' dimensions

- No issues were highlighted regarding the following dimensions: 'meaningful' , 'valuable', 'quality of information sufficient to support communication with others', 'possible to interpret information to support needs', and 'extent it is influenced by organisational beliefs, culture and expectations'

- The sourcing department raised the highest number of information quality related issues, followed by planners and buyers. International supply chain and finance had the least number of issues.

4.12. Combined findings

Findings are now summarised from the data collected from different business information stakeholders:

- All business information stakeholders involved in this organisation's ordering process for imports, experience information quality related issues, regardless of whether they operate at a strategic or operational level

- Information stakeholders experience different information quality related issues depending on their roles. Some stakeholders experience more issues than others

- Not all business stakeholders experience issues for the same information quality dimensions. This means that a dimension that is a problem for one group of business stakeholders is not necessarily a problem for the next group

- Information 'accuracy' or 'correct' information is the dimension with the highest number of information quality issues in this business process, followed by 'accessibility'. Every group of business stakeholder involved in this business process had issues to highlight about the 'accuracy' or 'correct' dimension. Similarly, with the exception of finance, every other group had issues to highlight about the 'accessibility' dimension

- For some information stakeholders, information quality issues experienced are a result of not following the business process and for others, they are a result of bugs in information systems.

While the syntactic aspect of information seems to have been sufficiently addressed by the systems, the semantic aspect had issues where stakeholders were not always able to attach sufficient meaning to the data. The pragmatic aspect of information quality was also an issue as stakeholders could not always use the information consumed to perform their tasks and therefore had to perform additional data manipulations. As for the social aspect, the information quality issues highlighted had a negative influence on the relationship between stakeholders.

CHAPTER FIVE - DISCUSSION

5.1. Introduction

This chapter is introduced with a summary of findings from data collected during interviews with business stakeholders to understand the chosen business process, as well as data collected using the semiotic information quality and categorical information quality matrix. These findings are then compared to each other before they are compared with literature.

5.2. Information quality findings

The table below summarises findings from the previous chapter.

Table 5.1: Summary of findings

Findings from the Business process understanding	Findings from a categorical information quality perspective	Findings from a semiotics information quality perspective
Most activities in this business process are managed in an automated way through sustainable information systems. However, a number of activities still rely heavily on excel spreadsheets.	Most information quality issues were related to the 'accuracy' dimension, followed by 'accessibility' and 'consistent representation'	The 'Correct' dimension was associated with the highest number of information quality issues, followed by 'Accessible' and 'non-redundant' dimensions
Every business stakeholder role involved in this business process experiences information quality related issues. Whenever information quality issues are experienced, the business process is directly affected.	No issues were highlighted regarding the 'objectivity', 'reputation', 'access security', 'relevancy', 'value-added', 'completeness' and 'concise representation' dimensions	No issues were highlighted regarding the following dimensions: 'meaningful' , 'valuable', 'quality of information sufficient to support communication with others', 'possible to interpret information to support needs', and 'extent it is influenced by organizational beliefs, culture and expectations'
Not all stakeholders agree on the causes of information quality issues experienced. Some believe that these issues are caused by business stakeholders not following the process and others believe that information quality issues are a result of bugs in information systems.	The sourcing department, followed by planners and buyers raised the highest numbers of information quality related issues	The sourcing department raised the highest number of information quality related issues, followed by planners and buyers. International supply chain and finance had the least number of issues.
	International supply chain department raised the lowest number of issues but it is also important to note that a lot of data needed to perform their job is not available in any automated system	

5.3. Comparison of information quality findings

This section compares all the findings summarised in the previous section by highlighting similarities and differences.

Several similarities can be extracted from the findings. This is particularly the case when comparing findings arrived at using the categorical information quality perspective with those arrived at using the semiotics information quality perspective. The biggest similarity with these two perspectives that the accuracy dimension, referred to as the 'correct' dimension in the semiotics information quality matrix, was related to the highest number of issues highlighted.

The fact that the number of issues highlighted by business stakeholders was ranked in descending order of sourcing department, followed by planners, then buyers in the third position, was also a big similarity in the findings.

The international supply chain department had a significantly lower number of issues for both perspectives, but this can probably be explained by the fact that a great deal of information used for their decision making is not available in computerised systems at the moment. This means that their biggest information quality issue is the lack of information.

Common findings across all three methods used to collect data were that all stakeholders involved in this business process did experience information quality related issues, as well as the manual use of excel spreadsheets to conduct certain activities in the business process being the cause of some information quality issues.

Many of the softer or less tangible dimensions, such as those in the social category of the semiotics framework, had no issues raised against them.

There were not too many differences in the data collected across the three methods used. The biggest difference was that stakeholders involved at different stages of the business process did not agree on the cause of information quality issues that they experienced. Opinions were divided between information quality issues being a result of bugs in information systems versus business stakeholders not following the process.

5.4. Information quality findings and the literature

This section compares the findings summarised in section 5.1 to the literature review in chapter 2.

One of the findings from the interviews conducted was that business stakeholders involved in the chosen business process did not agree with the causes of information quality issues

experienced. Depending on their role in the business process, some stakeholders felt that these issues were caused by bugs in information systems, therefore suggesting that information quality issues are an information technology problem, and others believed that the issues were a result of business stakeholders not following the business process. This is aligned with English (2010), summarised in table 2.7, when he states that Information quality is both a business problem and a system problem. In addition to that, Loshin (2011) in the virtuous cycle for information quality presented in figure 2.4 and summarised in table 2.8 makes it clear that information custodians and information managers cannot assess the quality of information in isolation from information consumers. This point emphasises the fact that information quality is everybody's business. The findings from the interviews highlight a gap in the organisation's understanding of information quality that this research could help bridge. In order to resolve information quality issues highlighted, business stakeholders will need to understand that these issues are both a business problem and a system problem, not just the one or the other.

Based on work by Graefe (2003) represented in figure 1.2, the quality of information provided to a consumer influences the utility of information in their decision making process. He suggests that Information consumers' decisions are influenced by how believable they consider the information to be. One of the issues highlighted by the buying department during the interviews was that they often get multiple Foreign exchange rates available for the same purchase order. These are the kind of issues that could result in buyers to stop trusting the information provided to them and thereby reduce the utility of information in their decision making process. This observation is also in line with the concept of information value chain which according to Berner (2012) aims to increase usefulness of information to the consumer and enabling them to make better decisions. Cases such us the multiple exchange rates or even the fact that planners do not always have access to planning data when needed, do not increase the usefulness of information, nor do they enable consumers to make better decisions.

Some business stakeholders highlighted that whenever they encountered information quality issues, activities and decisions in the business process were directly affected. For example, finance is not able to pay suppliers in time whenever supplier master data is not up to date. This is also aligned with the literature when Loshin (2011) stated that poor information quality affects the way in which business processes are intended to run. Conversely, poorly executed business processes result in poor quality information. For example, not following the correct process for receiving the order could result in incorrect shipment quantity declared. It is, therefore, important for business stakeholders to understand that there is a strong relationship between information quality and business processes.

After capturing the data collected from interviews using the categorical information quality matrix and the semiotics information quality matrix, it became very apparent that most of the issues raised were related to the accuracy or correctness dimension. This could easily suggest that information quality is mainly treated as an intrinsic concept in the organisation, independent of the context in which information is produced and consumed. However, there are some other pertinent information quality issues, related to other dimensions, which if ignored, could lead to serious implications. For example, issues where the quality of information is not sufficient to support relationships between stakeholders. This was the case with the quality of supplier master data which could result in suppliers not being paid in time. Such issues could lead to a termination of the relationship between the supplier and the organization. Similar to this example, is the case where there is no real-time information available to compare the planned against the actual margin for a supplier. This information is used in assessing the performance of a supplier. Therefore it is critical in maintaining the relationship between the supplier and the organization. This argument is also aligned with the literature when Strong et al., (1997) warn one that paying attention to accuracy alone does not correspond to information consumers' broader information quality concerns. When quality problems are reduced to just as errors in stored information, information quality professionals may not recognise and thus solve the most critical information quality problems in the organisation.

It can therefore be concluded that quality problems associated with the information as it flows through the organisational business processes are present in all quality categories similar to what is reported in the literature. However, as long as the social aspect is not specifically addressed quality problems will never be regarded with a shared understanding by the different information stakeholder groups.

CHAPTER SIX - CONCLUSION

6.1. Introduction

This chapter concludes the research. It starts with an overview of the different chapters of the thesis, then the limitations of the study, before it discusses possible further research.

6.2. Overview of chapters

This thesis has six chapters. The first chapter provides an overview of the research indicating the reason for conducting the study as well as its objectives. A statement of the research problem and research questions is given. A brief literature review is given to support the need for the research. A conceptual framework for the study, the research design, data collection, data analysis, delineation of research, and contribution of research form the rest of the chapter.

Chapter two is a discussion of information quality literature reviewed. It starts with acknowledging the difference between the terms ‘data’ and ‘information’, then introduces the concept of value chain, before extending it to information value chain. Different interpretations of information quality, information quality dimensions, information quality frameworks, information stakeholders, the view of information as a product as well as information as a service, information quality measurement, information quality management, and, finally, information quality management in the age of big data, form the rest of the literature review.

Chapter three discusses the research methodology followed in the study. The research philosophy, methodical choice, research strategies, time horizon of the study, data collection, data analysis, unit of analysis, unit of observation and research design are discussed.

Chapter four describes the business case study by firstly giving the background of its context and operation before the empirical data collected is analysed using two information quality frameworks discussed in chapter two.

In chapter five, the research findings are summarised, compared against each other and compared to the literature review.

Chapter six is the final chapter of this thesis, in which chapters are summarised; the research question is revisited, followed by a summary of the main contributions. The chapter is concluded by highlighting limitations of the study as well as giving directions for further research.

6.3. Research questions re-visited

The purpose of the research questions was to explore how different business stakeholders view information quality in an organisation's information value chain for a chosen business process, and what kind of information quality issues they experience. The findings from the research indicated that there no single view on the quality of information. Although all information stakeholders involved in the chosen business process experience information quality issues, they do not necessarily experience the same kind of issues, and the different dimensions of information quality do not have the same level of importance for all business stakeholders.

- **How do business stakeholders view the quality of information throughout the information value chain of a retail organisation?**

  The answer to this question is that information quality is complex and there is no single view shared by all stakeholders involved in an information value chain. While some stakeholders only experience one or two information quality related issues in a particular business process and, therefore, feel that the quality of information is good enough to perform their jobs, others experience a lot more issues and, therefore, do not share the same opinion. Findings from the business case, discussed in chapter four, highlighted that stakeholders also did not agree on causes of information quality issues experienced. Some felt that the issues were a result of stakeholders not following the business process and others felt that those issues are a result of bugs in information systems. In order to resolve information quality issues in the organisation, all stakeholders would need to understand that these issues are a business problem as well as a system problem, not just the one or the other.

  The following research sub questions were also addressed in the case study:

- **What kind of information quality issues do business stakeholders experience throughout the information value chain of a retail organisation?**
  Information stakeholders experience different information quality issues depending on their role. Chapter four of this study presented and categorised these issues in the multiple information quality dimensions using different frameworks. Some examples of issues included incorrectly captured data, ambiguity in information presented to consumers, Inaccessible information, issues that affect relationships between stakeholders and many more.

- **How are stakeholder decisions affected by information quality when producing or consuming information at the different touch points of the information value chain?**

  Findings from this study discussed in chapter five, confirmed that incorrect decisions are made by stakeholders when producing or consuming poor quality information

  Answers to the above questions were discussed in detail in chapter four and chapter five.

  The following section establishes the contribution of this research to the IS field.

6.4. Summary of the research contribution

Contributions of this research are theoretical, practical and methodological. The theoretical contribution is:

- A conceptual framework showing the relationships between different concepts discussed (Presented in chapter 2)
- An application of the value chain concept to information management (Discussed in chapter 2)
- Different views or interpretations of information quality (Discussed in chapter 2)
- A list of information quality frameworks and their related information quality categories and dimensions (Discussed in chapter 2)
- An association between information stakeholder roles and information quality categories (Discussed in chapter 2)
- A list of information quality management principles (Presented and discussed in chapter 2)
- A framework for information quality management (Presented in chapter 2)
- Recommendations for information quality management in the age of Big data (Discussed in chapter 2).

The methodological contribution is:

- An application of a research design framework (presented and discussed in chapter 3)
- A detail description and discussion of the case study (in chapter 4).

The practical contribution :

- Issues made visible in this study can be used to improve the organisation's information quality
- The frameworks used in this study can also be used to collect and understand information quality issues experienced in other business processes within the chosen organisation and in other organisations.

6.5. Further research

This research was more theoretical and conceptual because of the nature of the research problem. The importance of applying the concept of value chains to information management with the aim of improving information quality was highlighted.

Further research is required to explore how stakeholders in different line of businesses within the chosen organisation view information quality, to create visibility on information quality issues experienced and ultimately improve the entire organisation's information quality. A deeper investigation on dealing with information quality in the age of big data is also required.

REFERENCES

Al-Hakim, L. 2007. Information Quality Management: Theory and applications. Idea Group Publishing, London.

Batini, C & Scannapieca, M. 2006. Data quality: Concepts, methodologies and techniques. Springer, Berlin.

Berner, S. 2012. The value of managing information: How information management facilitates process improvement in corporations.

Boyce, C. & Neale, P. 2006. Conducting In-depth interviews: A Guide for Designing and Conducting In-depth Interviews for Evaluation Input. Available from: http://www.pathfind.org/site/DocServer/m_e_tool_series_indepth_interviews.pdf?docID=6301

Bryson, J.M. 2004. What to do when stakeholders matter: stakeholder identification and analysis techniques. Public Management Review, 6(1), 21-53.

Caballero, I., Verbo, E., Calero, C. & Piattini, M. 2007. A data quality measurement information model based on ISO/IEC 15939. Available from: http://mitiq.mit.edu/iciq/PDF/

Davidson, B., Lee, Y. & Wang, R. 2004. Developing data production maps. International *journal on healthcare technology and management 6(2), 87-103.*

Carr, W. & Kemmis, S. 1986. Becoming critical: Education, knowledge, and action research. Falmer, London.

Chen, M., Ebert, D., Hagen, H., Laramee, R., Van Liere, R., Ma, L., Ribarsky, W., Scheuermann, G. & Silver, D. 2009. Data, Information, and knowledge in Vizualization. IEEE Computer Graphics and Applications , 29(1):12-19.

DAMA international. 2010. The DAMA Guide to The Data Management Body of Knowledge (DAMA-DMBOK Guide), (1st ed.).

Dedeke, A.2000. A conceptual framework for developing quality measures for information systems. Proceedings of the fifth international conference on information quality.

De La Harpe, M. 2008. Organizational implications of data quality: A social perspective. PhD Thesis. Cape Peninsula University of Technology, South Africa.

De la Harpe, R. & Roode, D. 2004. An actor-network theory perspective on data quality in medical practices. Studies in Communication Sciences. 2:69-84.

Dunn, C., Cherrington, J. O. & Hollander, A. S. 2005. Enterprise information systems: a pattern-based approach, (3rd ed.)., McGraw-Hill Irwin.

English, L. 2010. Quality principles, processes and techniques applied to information quality management. Proceedings of the fourth MIT Information Quality Industry Symposium. Available from: http://mitiq.mit.edu/iqis/documents/cdoiqs_201077/papers/01_06_t2d.pdf

English, L. 2008. Grounding IQ management practices in sound quality management systems. INFORMATION IMPACT International, Inc.

English, L. 2006. The 14 points of Information Quality Transformation. DM Review Magazine.

English, L. 2002. Process management and information quality: How improving information production processes improves information (product) quality. Proceedings of the Seventh International Conference on Information Quality.

English, L. 1999.Improving data warehouse and business information Quality: methods for reducing cost and increasing profit. Wiley computer publishing, New York.

Eppler, M. J. & Wittig, D. 2000. Conceptualizing information quality: A Review of Information Quality Frameworks from the Last Ten Years. Proceedings of fifth International Conference on Information Quality.

Friedman, T. 2014. Data quality assurance in the age of big data: Why and how. Gartner, Inc.

Ganesan, E. 2013. Process Hierarchy and Granularity Definition in Enterprise Process Modeling. Requirements Networking Group. Available from: http://ssrn.com/abstract=1929715

Giannoccaro, A., Shanks, G. & Darke, P. 1999. Stakeholder Perceptions of Data Quality in a Data Warehouse Environment. Proceedings from the tenth Australasian conference on information systems.

Graefe, G. 2003. Incredible information on the Internet: Biased information provision and lack of credibility as a cause of insufficient information quality. In: M. Eppler & M.Helfert (Eds.). Proceedings of the Eighth International Conference on Information Quality (IQ 2003):133-146. MIT 2003. 7-9 November 2003.

Herzog, T., Scheuren, F. & Winkler, W. 2007. Data quality and record linkage techniques. Springer, New York.

Hughes, I. & Seymour-Rolls, K. 2000. Participatory Action Research: Getting the Job Done. Action Research E-Reports. Available from: http://www.fhs.usyd.edu.au/arow/arer/004.htm (Accessed 19 August 2010)

Jeffery, N. 2009. Stakeholder engagement: A road map towards meaningful engagement. Cranfield School of Management, U.K.: Doughty Centre for Corporate Responsibility: 1-46.

Jiang, L., Borgida, A., Topaloglou, T. & Mylopoulos, J. 2008. Data quality by design: A goal oriented approach. Proceedings from the twelfth International conference on information quality.

Jones, D. 2008. Essential data quality skills #1: Information Chain Management. Available from: http://dataqualitypro.squarespace.com/data-quality-home/essential-data-quality-skills-1-information-chain-management.html

Kahn, B., Strong, D. & Wang R. 2002. Information quality benchmarks: Product and service performance. Communications of the ACM 45 (4ve): 184–192.

Karr, A., Sanil, A., Sacks, J. & Elmagarmid, E. 2001. Affiliates workshop on data quality. National Institute of Statistical Sciences. Available from: www.niss.org/affiliates/dqworkshop/report/dq-report.pdf.

Klapinsky, R. & Morris, M. (n.d). Handbook for value chain research. Available from: http://sds.ukzn.ac.za/files/handbook_valuechainresearch.pdf; (Accessed 06 July 2013)

Knight, S. & Burn, J. 2005. Developing a framework for assessing information quality on the world wide web. *Informing Science Journal 8, 159–172*

Knight, S. 2011. The combined conceptual life-cycle model of information quality: part 1, an investigative framework*. International Journal of Information Quality, 2, 205–230.*

Knight, S. 2005. The combined conceptual life-cycle model of information quality: Part 1, an investigative framework. *International journal of information quality, 2(3), 205-230.*

Lam, K. (n, d). A study of business process reengineering. Available from: http://www.doc.ic.ac.uk/~nd/surprise_95/journal/vol2/tmkl/article2.html (Accesses 20 July 2010).

Laudon, J. & Laudon, K. 2006. Essentials of Business Information Systems. 7th ed, Pearson Prentice Hall.

Lee, Y. & Strong, D. 2003. Knowing-why about data processes and data quality. *Journal of Management Information Systems, 20(3), 13-39.*

Loshin, D. 2011. The Practitioner's Guide to Data Quality Improvement. Morgan Kaufmann Publishers, Burlington, MA.

Martin, M. 2005.Measuring and improving data quality part II: Measuring data quality. MPH, Clemson University

McCormick, D. & Onjala, J. 2007. Methodology for value chain analysis in ICT industry: frameworks for the study of Africa. University of Nairobi.

Mohanty, S., Jagadeesh, M. & Srivatsa, H. 2013. Big Data Imperatives: Enterprise Big Data Warehouse, BI Implementations and Analytics. Apress

Naumann, F. & Rolker, C. 2000. Assessment Methods for Information Quality Criteria. Proceedings of the fifth International Conference on Information Quality

Neely, M.P. 2005. The product approach to Data Quality and fitness for use: A framework for analysis. Proceedings of the tenth international conference on information quality.

Niemann, I. 2005. Strategic integrated communication implementation: Towards a South African conceptual model. PhD thesis. University of Pretoria, South Africa.

Nikolaou, A. 2009. Integrated information systems and transparency in business reporting. *International Journal of Disclosure and Governance 7, 216–226.*

Open Data Center Alliance, Inc. 2013. Open Data Center Alliance: Information as a Service Rev. 1.0. Available from: http://www.opendatacenteralliance.org/docs/Information_as_a_Service_Master_Usage_Model_Rev1.0.pdf.

Oz, E. 2006. Management information systems. (5th ed.). Boston: Thomson Course Technology

Porter E, M. 1986. Changing patterns of international competition. California management review.

Redman, T. 2001. Data quality: the field guide. Boston: digital press, 2001

Reeve, A. 2013. Managing Data in Motion: Data Integration Best Practice Techniques and Technologies. Morgan Kaufmann Publishers. Books24x7.

Runeson, Per. & Host, M. 2008. Guidelines for conducting and reporting case study research in software engineering. Empir Software Eng 14:131-164

Saunders, M. & Tosey, P. 2013. The layers of research design. Available from: Available from: http://www.academia.edu/4107831/The_Layers_of_Research_Design (Accessed on 08 December 2014)

Schlögl, C. 2005. Information and knowledge management: dimensions and approaches. Information Research. 10(4)

Schmarzo, B. 2013. Big Data: Understanding How Data Powers Big Business. Wiley, Hoboken.

Schwolow, S. & Jungfalk, M. 2011. Information value chain. Available from: http://www.informationvaluechain.com/wp-content/uploads/informationvalue-chain.pdf

Sebastian-Coleman, L. 2013. Measuring Data Quality for Ongoing Improvement: A Data Quality Assessment Framework. (1st ed.), Morgan Kaufmann.

Shanks, G. & Corbitt, B. 1999. Understanding Data Quality: Social and Cultural Aspects. Proceedings of the 10th Australasian Conference on Information Systems. Wellington, New Zealand

Strong, D.M., Lee, Y.W. & Wang, R.Y. 1997. Data quality in context. Communications of the ACM, 40(5):103-110.

Talburt, J. 2011. Entity resolution and information quality. Morgan Kaufmann Publishers.

University of Alberta. 2004. U of A management information value chain. Available from: http://www.uofaweb.ualberta.ca/vpfinanceais/pdf/AISSC_Sept-4_Management_Information_Value_Chain-03TkMC.pdf
(Accessed 22 April 2010)

Waddington, D. 2009. The sad state of data quality. Information Management Magazine. 19(8).

Walsham, G. 2006. Doing interpretive research. *European journal of Information systems 15, 320 - 330.*

Wang, R., Allen, T., Harris, W. & Madnick, S. 2003. An information product approach for total information awareness. Proceedings of the IEEE aerospace conference, Big Sky, MT.

Wang, R., Pipino, L.& Lee, Y. 2002. Data Quality Assessment. Communications of the ACM 45(4), 211-218.

Yin, R. 2003. Case Study Research: Design and Methods. Sage, Thousand Oaks, CA, 2003b, (3rd ed.).

Yläranta, M. 2006. Between two worlds - Stakeholder management in a knowledge intensive governmental organisation. Publications of the Turku school of economics. A-7:2006.

APPENDIX: Initial Information Quality Interviews

The researcher conducted initial interviews with the purpose of understanding the chosen business unit in the organization as well as its business processes. Another objective of the initial interviews was to get a high level view on the type of information quality issues that the business unit faces. Two respondents with in-depth understanding of business information and business processes took part in these initial interviews. As a result of the initial interviews, a particular business process was chosen and subsequently used as business case for the study. Data collected from the business case's interviews was presented and analysed in chapter 4. This also formed the basis for all the other interviews conducted with the different information stakeholders from the different departments.

.

- **Interview 1:**

Questions	Respondent 1
What is your role in the company?	Process expert - Business Analyst with an in-depth understanding of current business processes, and a conceptual view of future processes. Focuses on Business improvement.
How do you use information in your work?	Everything she does uses information. Uses information to validate that business processes are working and that they are working properly. She uses information to justify reason for change in a process or to validate / prove that the process is working. Uses information at all level – High level and details If the information is incorrect, the business wastes lot of money. Respondent Thinks that there isn't enough supply chain data in the business. Business doesn't wait for IT to catch up therefore they manage their business on Excel or paper.
Do you use paper based information?	All Information that respondent 1 uses has been computerised. Her entire life has changed, everything is computerised. However they extract information from systems and manipulate it in some cases (e.g.: To perform what-if analysis, etc).
How much information do you use?	Lots of information

Questions	Respondent 1
How important is the quality of the information to you/organisation? (Why?).	Massively important. More important now than it has ever been before. The business has grown a lot and the amount of information is massive, therefore there is no way it can be managed manually. As you grow complexity in the business you don't have the time to look at information manually. Wrong information will results in wrong actions / decisions and that will impact the growth of your market and your bottom line directly.
What do you (or other employees) do to increase the quality of the information? (If nothing, Why? Is everything perfect?).	Respondent 1 adds value to information for analysis purposes (What-if analysis, draw graphs, Construct excel reports etc.,)
What are the most common information quality issues that they are confronted with? Example?	Whenever information comes out of a reputable environment, she finds no reason to think it is of a lesser quality. Respondent 1 believes that bad information is a result of a human hands having done something wrong (users / data capturer, etc). It's normally compliance or lack of compliance to Business processes that causes this. It's like a relay race, where the baton drops when it's passed on. People must fix causes. Understanding the end in mind when designing business processes. Design business processes with measurement and information in mind (Make sure that the information will be recorded in a system when designing business processes)
What are the obstacles that contribute to the information quality problem in your opinion? (elaborate)	• Management i.e. Insufficient leadership knowledge • Management focus (following up and looking into details) • Lack of consequence for poor performance • Staff turnover is too high in retail at the moment – people don't have the knowledge • Poor leadership • Complexity in Business processes – therefore you need specialists
Where do they get their information from? (Which Sources / Systems)	Because of the nature of respondent 1's job, 99% of information that they need is for analytical purposes, and they get it from BI reporting systems, mainly the cubes. Unfortunately respondent 1 cannot run scripts, so they rely very heavily on the standard BI environment. A number of their colleagues extract information from Warehousing systems directly (Triceps, Allpoints) to do both analysis as well as operational support.

Questions	Respondent 1
Do you ever need to extract the same information from multiple sources? If yes, why do you need to do that?	I do sometimes validate the data warehouse (BI) information against Retek Merchandise System (RMS) information but that is by exception and when I need to see the actual transaction dates of the data or where I need to drill to a level of information that the BI cubes do not display. e.g.: This week I was looking at Chain Intake/GRN's in week buckets from a cube and some of the values were slightly misleading. To understand the behaviour of the Warehouse I logged into RMS to check on the information at a more granular level. I may have been able to run reports from the data warehouse but I find many of the reports very slow and as I am usually impatient I find it easier to troll the RMS information
If you ask the same question from 2 different systems, do you always get the same answers? Give examples.	In certain instances the answer is the same. Where this happens it is because both systems do not manipulate the data. We do have differences with transactional data between planning systems, merchandising system and the data warehouse. An example would be PO delivery dates. In planning systems a planned delivery that is not received rolls to a new week. In the merchandising system and the data warehouse nothing "rolls". So, you may see a PO in the merchandising system with delivery intent of week 40 but in the planning systems, week 40 is history and the order reflects against week 41. The same would occur in a size profiling analytical system. It takes RMS's sales history by sku/location day or week and based on statistical logic will discard some history as statistically invalid for the purpose of proposing a future size profile – nothing wrong with this logic of cleansing data.
Is there anything that you believe could add value to my research.	Respondent 1 suggested that the researcher conduct other Interviews with members of the different business unit such as supply chain, for which a lot of data is not captured into information systems

- **Interview 2:**

Questions	Respondent 2
What is your role in the company?	Planning manager & Business Improvement manager
How do you use information in your work?	As a Planning Manager, respondent 2 Uses information to perform planning function in the life-cycle of product (i.e. Plan product, track products all the way through to stores and how they perform). Respondent Uses information for planning products and making trading decisions and therefore quality is crucial. For example: Items incorrectly assigned to Short sleeves instead of long Sleeves in the middle of winter will result in a disaster. This will have direct impact on profitability. Retailers who don't have information struggle. As a Business improvement manager, respondent 2 uses information to investigate what is wrong with the business process, in order to improve it. Therefore they need to trust the Quality of information.
Do you use paper based information?	Only paper based information they use are printed reports.
How much information do you use?	Lots of information
How important is the quality of the information to you/organisation? (Why?).	Bad information cost you money (in lost sales, margin, market share, etc.)
What do you (or other employees) do to increase the quality of the information? (If nothing, Why? Is everything perfect?).	As a Planning manager, respondent 2 extracts information into spreadsheets, manipulates it in order to make different trading decisions, e.g.: KPIs, Product benchmarking, etc. As a Business improvement manager, respondent 2 needs to view information differently sometimes and go back to IT with new requirement. (change in the presentation of the information)
What are the most common information quality issues that they are confronted with? Example?	Information not on time is a huge one because sometimes information is delayed between systems. Incorrect calculation / accumulation & aggregation of information is also a big issue. Information is generally correct at the most granular level.
What are the obstacles that contribute to the information quality problem in your opinion? (elaborate)	1. Delays through the IT infrastructure 2. Complexity of the retail business
Where do they get their information from? (Which Sources / Systems)	BI & Planning systems JDA Assortment

Questions	Respondent 2
Do you ever need to extract the same information from multiple sources? If yes, why do you need to do that?	1. Yes, but that mainly because of different levels that systems keep information at and because sometimes you need to view your information with other information that you can't find in the same system. 2. Different systems get updated at different times and therefore causing differences (timing issues). That causes people to extract same info from different sources to understand what's happening.
Do you believe that value is added to information as it flows through different systems within your organisation? If yes, Please Give an example.	Yes, value is added to the information. Extra calculation done, etc. Transactional systems will tell you that u sold 10 blouses, but BI tell you that u sold 10 blouses therefore growth on last year is …
If you ask the same question from 2 different systems, do you always get the same answers? Give examples.	Generally because of timing and aggregations, u don't get the same answers.
Is there anything that you believe could add value to my research?	Make the research practical. An information quality issue can have serious ramifications. Includes these ramifications / impacts into the thesis. Make it more relevant to Retail.
Do you sometimes have to remind people about the importance of the information quality? (Who? Why? How?)	Yes. People don't always appreciate and understand the impact of the wrong decisions that has been made from wrong information
Does anyone have to remind you about the importance of Information Quality?	No, The lower down you go the higher the focus on quality.
What do you do when you are confronted with a problem related to the quality of the information? Example? (aware of potential impact?)	Respondent 2 Tries to verify the information by checking it with another system, then logs the issue with IT service centre. They know what system is wrong based on their understanding of Business process & trends. E.g.: You sell R1, 000,000 every week, and one week the system says you have sold R 4,000,000 and yet it's not Christmas.
Do you record your experiences/actions/outcomes relating to information quality problems?	Incidents are recorded on an incident management system (with incident reports)
Do you share your experiences relating to information quality problems with others (why (not)?)	Based on their experience, respondent 2 shares their knowledge with Junior People in the team on how to recognise information issues.