CHAPTER 1 INTRODUCTION

1.1 Motivation for the study

Data-driven decision support systems are very expensive to develop. This is due to various factors, including the transformation of different data sources to a single platform and a high level of managerial involvement. The magnitude of the data sources involved requires large capacity hardware resources that are expensive. Another cost factor is that due to the nature of the processes simulated in the data-driven decision support system, involvement of senior management in the problem definition is essential. Markus (2000:43) argues that the implementation of enterprise systems in large corporations have been known to cost over $500 million and because systems are much more tightly integrated than before, failure of one system will have extremely negative consequences for other systems in the organisation.

Data warehouses are typical data-driven decision support systems. Data warehouses integrate various data sources to supply management information in organisations. Many data warehouse projects take longer than originally planned and cost much more than initially budgeted for. The Cutter Consortium reported that 41% of practitioners surveyed have experienced data warehouse failures (Anonymous, 2003:1). Inmon (reported in Ferranti, 1998:1) argues that companies make costly mistakes that cause delays. However, the benefits of successful data warehouses are so significant that the academic research community should search for methods to improve the success rate of data warehouse projects.

The author of this thesis was drawn to systems thinking as a basis for improving data warehouse quality because of the holistic nature of systems thinking. Kevin Strange of Gartner wrote (2001:1): “With respect to the analytical (business intelligence) side of customer relationship management, at least 65% of the efforts are implemented in an unintegrated fashion, based on a function (different efforts by different departments), rather than on a more strategic initiative – the sum is larger than the parts.” Many authors (Mimno, 2001; Eckerson, 2003) argue that business objectives should be central to data warehouse planning and development. This is congruent with the systems approach proposed by Churchman (1968). He advocates that subsystems should work together to achieve the objectives of the system and the

objectives of the subsystem should relate to that of the system. A question worth investigating is whether systems thinking can point practitioners to more successful data warehouse development practices.

This thesis is such a research initiative. The aim of the research is to develop a framework for the explicit use of specific systems thinking methodologies in data warehousing practices. It is assumed that data warehousing practitioners do not know systems thinking methodologies (a valid assumption according to the case study data reported in chapter 5). From data warehousing literature (Kimball *et al.,* 1998; Inmon, 1996), it is clear that the practices of data warehousing professionals can be mapped to systems thinking methodologies. The researcher decided to make this mapping explicit in order to propose methods, in the form of a framework, to improve data warehouse quality.

The development of the framework is seen as a two part process. The first part is to explore current data warehousing practices according to systems thinking methodologies. The second part is to make this mapping explicit in terms of a framework.

1.2 Concepts central to this thesis

Systems thinking and data warehousing have been introduced in section 1.1 as concepts central to the development of the researcher's research presented in this thesis. A short summary of these concepts will suffice for the time being, but they will be dealt with in detail in chapter 3 and chapter 4.

1.2.1 Systems thinking

The term methodology refers to methods for exploring and gaining knowledge about systems. Systems thinking emerged in reaction to reductionism, when Von Bertalanffy (1968:51) advocated an interdisciplinary approach to widen the scope when studying problem situations. A system is a set of interrelated entities, of which no subset is unrelated to any other subset (Kramer & De Smit, 1977:13) and has properties that do not exist in the parts but are found in the whole (Weinberg, 1975:60). Churchman (1968:29) describes systems in terms of their objectives,

environment, resources, components and their management. He argues that a specific system can be identified by its objectives. Different ontological views of systems, which we call methodologies, emerged, namely hard systems, soft systems, critical systems and recently disclosive systems thinking.

Hard systems thinking is a term used by Checkland (1981) as an alternative to "soft systems". From a hard systems thinking perspective social systems are treated like scientific problems. A system is viewed as a hierarchically organised set of elements. When one understands the components of the system, one is able to understand the system as a whole. A system is seen as a true representation of reality. Information systems development, according to a hard approach, is seen as a technical project which can be done outside the context of the environment.

Soft systems thinking adopts a more holistic approach to systems properties when declaring that a system has properties that do not exist in its parts. A system is viewed as a person's perception of the real world. Different views enhance the understanding of the problem situation. User satisfaction is more important than requirements conformation in information systems development. Checkland (1981, 1999) developed a methodology (set of methods) to implement soft systems thinking in solving problems.

Critical systems thinkers believe that the world is not fundamentally harmonious. Therefore, to understand, explain and make possible changes, one must think in terms of contradictions. Different perceptions can be seen as expressions of irreconcilable conflict and power struggle between management and workers, or systems developers and users (Flood & Jackson, 1991a:83). Intervention is central to practising critical systems.

Disclosive systems thinking was introduced by Strijbos (2000) to address the responsibility of people for particular developments. He accentuates the norms for action taken by agents and argues that ethics should be part of the chosen systems methodology. Strijbos warns that ethics is not part of hard, soft, or critical systems thinking. He states that "disclosive systems thinking (and the systems ethics entailed in it) proceed from the normative view that the various systems receive their meaning from the pre-given reality and order of which these systems are a part" (Strijbos, 2000:168). Disclosive systems thinking aims to disclose this intrinsic normativity in order to enrich human life and culture.

1.2.2 Data-driven decision support systems

Inmon (1996:33) defines a data warehouse as a "subject oriented integrated, non-volatile, and time variant collection of data in support of management decisions." Kimball *et al.* (1998:19) simply define a data warehouse as "The queryable source of data in the enterprise." These authors are most influential in data warehousing design methodologies. They differ on various concepts in data warehousing, one of which is the development lifecycle of a data warehouse. Inmon (1996:24) advocates a lifecycle that he calls the CLDS (reverse of SDLC: systems development lifecycle) with the following phases:

1. Implement data warehouse
2. Integrate data
3. Test for bias
4. Program against data
5. Design DSS system
6. Analyse results
7. Understand requirements

This is a data-driven lifecycle methodology. Kimball *et al.* (1998) advocate the use of a requirements-driven lifecycle methodology. Their methodology begins with a data warehouse readiness test. Then user requirements are gathered, followed by modelling, data staging, end-user application design, and maintenance.

Different authors identify success factors in data warehouse design. Ferranti (1998) quotes Inmon: "Building data marts before developing a data warehouse can be one of your biggest mistakes." Mimno (2001) argues that the most important success factor is to make your data warehouse business-driven. He argues that a technology-driven approach is much more likely to fail than a business-driven approach.

1.2.3 Relationship between systems thinking and data-driven decision support systems

Many data warehousing authors such as Kimball *et al.* (1998) and Mallach (2000) advocate user involvement, the inclusion of business sponsors, and the involvement of top management, to increase the success rate of data warehouse systems. A very strong implicit use of soft systems methods surfaces when one examines the

lifecycle of a data warehouse and the recipes for data warehousing success. Markus (2000:44) argues that the development process of business-driven systems, such as data warehouses, looks more like a large-scale organisational development or change management project than it looks like a traditional information systems (IS) project.

One might argue that a technology-driven approach is a hard systems approach, and a business-driven approach is a soft, critical, or disclosive systems thinking approach depending on the characteristics of business objectives. The aim of this research initiative is to apply systems thinking ideas on data warehousing in order to improve data warehousing quality. The first step however, is to understand current practices of data warehousing practitioners from a systems thinking methodology point of view.

1.3 Structure of the thesis

This thesis explores the relationship between philosophy, methodology, and practice, indicated in Figure 1.1. The philosophy layer represents general thinking based on ontological and epistemological assumptions. The works of individual philosophers as well as ideologies are represented by the "philosophy" section of this structure. "Methodology" is seen as general procedures used to explore reality. Section 3.1 explores the definition of "methodology" further. "Practices" represent activities of people in the performance of their daily work. It can be seen both as individual practices as well as common practices typically used by people to perform a specific task.

This thesis aims to demonstrate that the selection of practices followed in a problem situation relates to the methodological and philosophical underpinnings of those practices. In this thesis the structure of philosophy, methodology, and practice is applied independently to two problem situations. Firstly, it is used to support the selection of an unorthodox research methodology used in the empirical part of the research presented in this thesis. Research methodology is presented in chapter two according to this structure. The process regarding the selection of a specific research plan is presented as an application of this structure.

Secondly, it forms the core of the main argument of the thesis that motivations behind practices in data driven decision support systems development are rooted in

specific systems thinking methodologies, which are in turn rooted in specific philosophical ideas. This statement leads to the development of a framework for the use of a specific systems thinking methodology in data driven decision support development. Chapter 3 presents systems thinking according to this structure and chapters 5 and 6 aim to relate data driven decision support development practices to their methodological and philosophical underpinnings.

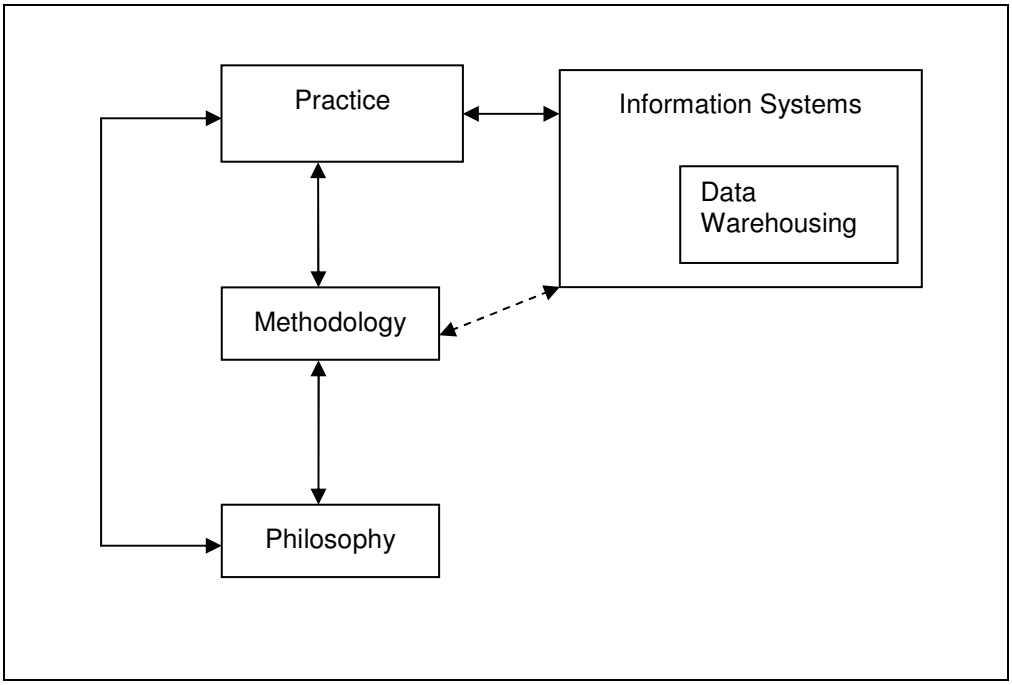

Figure 1.1 The relationship between philosophy, methodology, and practice

1.4 Research objectives

The main objective of this study is to transform the implicit relationship between systems thinking methodologies and data-driven decision support system development practices as reflected in literature and practice, to an explicit relationship. A framework for the use of specific systems thinking methodologies in data-driven decision support system development practices will be set up to encourage the explicit use of practices other than hard systems practices in data-oriented DSS development.

To reach this objective the following sub-objectives need to be achieved:

1. The first sub-objective is to understand data-driven decision support system development practices from a systems thinking point of view. This is done in

order to relate the final framework to the current practices of data-driven decision support system development practitioners. Three aspects need to be investigated to reach this objective:

a. The researcher needs to study systems thinking; this is done by means of a literature study that investigates the relationships between philosophy, methodology and practice of systems thinking and the application thereof in information system development.
b. The researcher needs to study data-driven decision support systems; this is done by means of a literature study on data warehousing and the success factors of data warehousing projects.
c. The researcher needs to explore the systems thinking nature of current practices of data warehousing professionals, even without the professionals being knowledgeable of systems thinking methodologies. This is done by means of case study research.

2. The second sub-objective is to design a framework that models specific systems thinking methodologies in the various stages of the development lifecycle of data-oriented decision support systems explicitly. The framework will be presented according to the systems approach introduced by Churchman (1968).

1.5 Limitations of the study

Data warehouses are investigated as representative of data-driven decision support systems. Mallach (2000:143) describes data warehouses as the primary data-driven decision support system used today. Therefore it is reasonable to limit this study to data warehouses to the exclusion of other data-driven decision support systems. The research cannot however automatically be applied to all data-driven decision support systems.

It was discovered during the initial approaches to the case study organisations that the methodologies of Inmon (1996) and Kimball *et al*. (1998) are dominant in the data warehousing industry in South Africa. This study therefore focusses on these generic methodologies to the exclusion of other specific information systems development methodologies.

Since data warehousing practices rather than data warehousing usage are the focus of the thesis, interviews were conducted with data warehousing practitioners rather than with end-users. The role of the end-users in the final framework is based on reports of the data warehouse team members interviewed.

Acceptance testing of the framework needs to be done through publication in journals and data warehousing specific literature and is viewed as further research.

1.6 Methodology

To reach the research objective, the data warehouse development process must be investigated and the implicit use of specific systems methodologies must be identified. The nature of the study forces the research to focus on techniques that facilitate field studies where the researcher is an active part of the environment. The researcher needs to assess the underlying motivations for the actions of the development team. It is important to understand whether motivations of actions are rooted in hard, soft, critical or disclosive systems thinking. The data warehouse development effort cannot be separated from the business strategy of the organisation. It is therefore required that the practices of the data warehouse development team are investigated in the context of the organisation as a whole. It is clear that interpretive methods must be used as opposed to positivistic methods.

In interpretive methods, the researcher serves as an instrument of observation (Lee, 1999:21). This is in contrast to positivistic methods where the researcher is an objective onlooker who does not influence the situation at all. Lee (1999:22) further argues that positivism and interpretivism are incomplete for information systems (IS) research because the researcher does not influence and is not influenced by the research environment. He argues that IS research should take on another form of research namely “critical social theory (CST)”. CST researchers believe that no researcher can simply be an onlooker, but that the researcher influences and is influenced by the social and technological systems he/she is studying. The researchers must also play a role in the emancipation of the actors in the research environment from unjust and inequitable conditions. Action research introduced by Lewin (1947) is an example of CST research methods.

Three aspects differentiate this study from typical interpretive research case studies. Firstly, in interpretive case studies, theory is typically generated from data or observation. Since the data warehousing practitioners are not knowledgeable on systems thinking methodologies, theory generated from their practices will not reflect systems thinking methodologies. Secondly, in order to identify what to investigate, the researcher needs to create a preliminary mapping between systems thinking and data warehousing practices. This is against the nature of typical interpretive research as described in chapter 2. Thirdly, the aim of the study is to eventually change (improve) the practices of data warehousing professionals, that is, to intervene. From this intervention aspect, one might argue that this research should be classified as critical social theory.

However, the study differs from typical critical methods, such as action research, in that the praxis or intervention is not done as part of the research process. A framework is not proposed to be used to implement a data warehouse and then improved according to the results of the implementation, followed by several repetitions of this cycle. The framework developed in this study is a result of case study data analysis.

This study can be compared with the diagnostic part of action research. Since the hermeneutics of interpretive research is very helpful, the researcher chose to declare this a pluralistic approach incorporating both interpretive and critical methods. Chapter 2 includes a detailed discussion of this argument.

The aim of the case studies is to understand the practices of the data warehousing professionals from systems thinking methodology point of view. Since the case study evidence is analysed through pattern matching (described by Yin, 1994:106), a preliminary mapping had to be developed prior to the first case study. Semi-structured interviews were conducted to gather data, and the answers to questions were mapped to specific systems thinking methodologies according to the preliminary mapping. This mapping was designed from data warehousing literature, combined with systems thinking literature and forms the basis for the case study data analysis. Questions were posed to data warehousing team members on six major data warehousing aspects: data warehouse adoption, data warehouse development methodology, requirements collection, data modelling, data staging (including data quality) and end-user applications. For each question, a typical answer was formulated in terms of each of the systems thinking methodologies investigated. No

other mapping between systems thinking methodologies and data warehousing practices could be found in literature.

The data collected from the case studies was combined with the knowledge gained from the two literature studies to develop the framework for the explicit use of systems thinking methodologies in data warehousing practices. This framework is presented in chapter 6.

1.7 Chapterisation

Chapter 2 gives a discussion on the research methodology used in the context of information systems research. The chapter is organised in terms of the philosophy, methodology, and practice model used throughout this thesis.

Chapter 3 reports on a literature study on systems thinking, in terms of philosophy, methodology and practice. It also investigates the application of systems methodology and practice in information systems development.

Chapter 4 reports on a literature study on data warehousing. This chapter includes a discussion on success factors in data warehousing practices.

Chapter 5 contains the case study reports. It gives a detailed description of the data analysis done in the study. A preliminary mapping of systems thinking methodology on data warehousing practices used in the data analysis of the case study data is discussed. It gives interpreted results of three case studies conducted to study data warehousing practices from a systems thinking point of view. The chapter concludes with an argument for the use of specific systems thinking methodologies based on the case study data analysis results.

Chapter 6 describes the final framework for the explicit use of specific system thinking methodologies in data warehousing practices. The framework is presented according to the systems approach introduced by Churchman (1968). The framework constitutes the conclusion of this thesis. The chapter also includes a summary of the research described in this thesis as well as a critical evaluation of the scientific contribution of the reported research.