International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:5, No:6, 2011

Improved Data Warehousing: Lessons Learnt from the Systems Approach

Roelien Goede

***Abstract*—**Data warehousing success is not high enough. User dissatisfaction and failure to adhere to time frames and budgets are too common. Most traditional information systems practices are rooted in hard systems thinking. Today, the great systems thinkers are forgotten by information systems developers. A data warehouse is still a system and it is worth investigating whether systems thinkers such as Churchman can enhance our practices today. This paper investigates data warehouse development practices from a systems thinking perspective. An empirical investigation is done in order to understand the everyday practices of data warehousing professionals from a systems perspective. The paper presents a model for the application of Churchman's systems approach in data warehouse development.

***Keywords*—**Data warehouse development, Information systems development, Interpretive case study, Systems thinking

I. Introduction

THE aim of this paper is to discuss how the systems approach can enrich data warehouse development. The systems approach of Churchman is often used to gain a better understanding of a problem environment [1]. Data warehousing is still the main data source for Business intelligence [2]. The success rate of data warehousing projects is not satisfactory [2]. This paper uses the systems approach to enhance the understanding of data warehousing.

The paper starts with a motivation for the study. A short discussion of data warehouse development follows in section III. The systems approach is then briefly discussed in section IV. An empirical study was done to investigate the systems thinking orientation of data warehouse practitioners. The research process is described in section V. Section VI gives guidelines of how the systems approach could enrich data warehouse development.

II. Why a systems approach to data warehousing?

Most data warehouse professionals are younger people who were educated in the late 1990's and early 2000's. During this period the ideas of Churchman and other systems thinkers were taken for granted and not explicitly taught in information systems training. Not a single participant of the 40 people interviewed for this study heard about systems thinking and systems thinking methodologies before this study. The ideas of Churchman and other systems thinkers are applicable to data warehousing issues today. An example is source data ownership. Source data quality remain one of the most problematic areas in data warehousing design, since the owner of the source systems loyalty lies at the operational system. The source system owners in the organizations interviewed do not view themselves as part of the data warehouse system, neither do the data warehouse team. If the boundary of the data warehouse is viewed differently from a systems point of view these source system owners should be viewed as part of the system.

Another example is Churchman's view of the components of the system, he argues that the departmental division of organizations in terms of sales, marketing, procurement etc is incorrect and that organizations should be seen and managed according to business processes. This is directly in line with Kimball's views for data marts in a data warehouse.

This paper aims to highlight some aspects of systems thinking and the systems approach that may be used to enhance data warehousing development.

R Goede is with the Vaal Campus of the Nort-West University, Vanderbijpark, 1935, RSA (phone: +27 169103276; e-mail: roelien.goede@nwu.ac.za).

III. Data warehousing

A. Definitions

Data warehouses are examples of decision support systems (DSS). A DSS can be defined as a "computer-based information system whose primary purpose is to provide knowledge workers with information on which to base informed decisions." [3]. DSS can be divided into data-oriented DSS, model-oriented DSS and process-oriented DSS. A data-oriented DSS uses data base systems as source of the decision support, in contrast to a model-oriented DSS which uses mathematical models to support business decisions and a process-oriented DSS which simulates human decision making processes [3]. Data warehouses are the primary example of data-oriented DSS today.

Data warehouses are also known as online analytical processing (OLAP) systems because they serve managers and knowledge workers in the field of data analysis and decision making.

Online transaction processing (OLTP) systems, or operational systems, are those information systems that support the daily processing that an organization does. OLTP systems' main purpose is to capture information about the economic activities of an organization. One might argue that the purpose of OLTP systems is to get data into computers, whereas the purpose of data warehouses is to get data or information out of computers.

Reference [4] describe the differences between data warehouses and OLTP systems

A literature study by [5] indicated two main authors in the field of data warehousing, namely William Inmon, who is known as the father of data warehousing, and Ralph Kimball. Their approaches to certain aspects of data warehousing differ greatly. Industry practitioners are aware of these authors and their differences. Practitioners choose to follow either an Inmon approach, or a Kimball approach. Other data warehousing literature can easily be labeled as more towards Inmon's, or more towards Kimball's ideas. Some of these differences will be highlighted in this section.

Reference [6] defines a data warehouse as a subject oriented integrated, non-volatile, and time variant collection of data in support of management decisions. Reference [7] explains each of the parts of this definition:

"*Subject oriented:* A data warehouse is organised around the key subjects (or high level entities) of the enterprise. Major subjects may include customers, patients, students and products. *Integrated:* The data housed in the data warehouse is defined using consistent naming conventions, formats, encoding structures, and related characteristics. *Time-variant:* Data in the data warehouse contains a time dimension so that it may be used as a historical record of the business. *Non-volatile:* Data in the data warehouse is loaded and refreshed from operational systems, but cannot be updated by end-users."

Reference [8] simply defines a data warehouse as "the queryable source of data in the enterprise."

B. Development methodologies

Inmon 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. Analyze results; 7. Understand requirements [6]. This is a data-driven lifecycle methodology. Reference [8] advocates the use of a requirements-driven lifecycle methodology. His methodology begins with a data warehouse readiness test, where after user requirements are gathered, followed by modeling, data staging, end-user application design, and maintenance. The aim of this section is to give the reader background knowledge on data warehousing without focusing on different strategies.

The aim of the data warehouse is to give end-users (mostly managers) easy access to data in the organization. In order to do this it is necessary to capture everyday operational data from the operational systems of the organization. Operational systems are transactional systems, for example point of sale systems that are designed around relational databases, which form the source systems of the data warehouse. The data from the source systems go through a process called data staging to the presentation servers [8]. Data staging involves four very important actions. Firstly, the data is extracted from the source systems. The data required for the data warehouse is usually distributed in various different source systems with different file formats running on different hardware and operating system platforms. Secondly, the data is transformed to the data warehouse format. Errors in the data and inconsistencies are removed during this phase. Thirdly, the data is loaded into data marts in the presentation server. The final task of the data staging area is to schedule this process.

Data access methods differ greatly between operational system and data warehouses. In operational systems fixed access methods are pre-built as standardized reports. These users use the data in a predetermined way. In data warehouses very few standardized reports are written. These users use browsers and ad hoc queries to access the data. Data in the data warehouse cannot be altered by the end-users, because of the historical nature of the data. It is possible however to add some of the report outputs of the end-users into data marts to enhance the data warehouse's functionality

IV. The Systems Approach

A. The emergence of systems thinking

Systems thinking is proposed as a method to overcome the shortcomings of the traditional scientific approach, it is necessary to briefly discuss the traditional scientific approach.

The Greek philosophers, Plato and Aristotle, developed the art of rational thinking, which forms the basis of scientific knowledge. Science is a way of acquiring publicly testable knowledge of the world. This knowledge is generally gained from rational thought combined with experience. The experience is gained from deliberately designed repeatable experiments. These experiments are designed to enable the scientist to formulate laws that govern the regularities in the universe. These laws are expressed mathematically. Three key aspects of the scientific method are reductionism, repeatability and refutation. Reductionism is the basis for removing complexity from problems. Descartes' second rule for "properly conducting one's reason", which is central to scientific problem solving, i.e. dividing up problems into separate parts, assumes that this division will not distort the phenomenon being studied [9]. This implies that components of the whole behave the same when studied separately as when they are part of the whole. Although this approach is reasonable for many physical phenomena in the world, it is very difficult to apply to problems in a more complex social environment.

Reference [10] defines reductionism as a doctrine that maintains that all objects and events, as well as their properties, and our experience and knowledge of them, are made up of ultimate elements, indivisible parts. All positivistic scientists identify something to form the basis element of their subject. Physical scientists believe that everything is made up of atoms; biologists believe that cells are the basic elements of life. Even Freud reduced personality to basic elements, i.e. id, ego, and superego.

In reaction to reductionism, [10] defines expansionism as a doctrine that maintains that all objects, events, and

experiences of them, are part of larger wholes. It does not deny that they have parts, but focuses on the wholes of which they are parts. During the 1940's the focus in philosophy shifted away from particles to symbols and later to languages. The context of the word in a whole sentence or phrase, is key to the understanding of that word.

Reference [9] discusses three problem areas of science: complexity, social science and management. Our knowledge is categorized into subject areas, to which we are so used to, that we have difficulty seeing the unity that underlines the diversity. This is done to help us simplify our world in order to make sense of reality, because of our limited ability to grasp the whole. Although most problems in physics can be explained with a manageable number of variables, which can be isolated in experimentation, it is very difficult for the biologist to do the same. When we examine social science in social reality, we find not only a large number of variables, but we are confronted with the question of value-free sociology. We are confronted with the question of whether the observer is able to stay objective, or whether he or she will participate subjectively in the organization. It is very difficult to design repeatable experiments in the social environment, owing to the unpredictability of social happenings. Managers often see their work as practice rather than science. Operational research and management science developed certain strategies to handle specific types of managerial problems (e.g. linear programming problems), by building models that represent reality. However, it is extremely difficult to estimate how accurately reality is represented by a specific model. There are countless situations in the everyday activities of a manager for which it is not possible to create models.

Reference [9] stresses that the aim of systems thinking is to tackle problems of irreducible complexity by thinking in wholes, rather than overthrowing the tradition of science.

When [11] declares: "A system is a way of looking at the world", he attempts to open up people's minds. He wants us to realize that people view things differently according to each one's own experience and point of view. Reference [11] further states that it is the purpose of the system that gives it its right of existence.

The systems approach considers the system as a whole, consisting of interdependent elements [12]. The specific arrangement of the parts of a system is significant. The environment and the interaction of the system with its environment cannot be ignored.

B. Definition of a system

Reference [10] defines a system as "a set of two interrelated elements of any kind; for example, concepts (as in the number system), objects (as in a telephone system or human body), or people (as in a social system)." The system is not indivisible but must be seen as a whole that can be divided into parts. A system is always more than the sum of its parts. A system's emergent properties are those properties that do not exist in the parts but are found in the whole [11]. A system also forms part of a larger whole or system.

C. The systems approach

Reference [12] developed a systems approach to address problem situations holistically. His work influenced many systems thinkers, such as Checkland and Jackson. Reference [12] declares that: "Systems are made up of sets of components that work together for the overall objective of the whole." He discusses five characteristics of a system, namely the total system objectives, the system's environment, the resources of the system, the components of the system, and the management of the system. If we analyze a situation using these characteristics, we follow what Churchman calls "the systems approach".

1) The total system objectives

When studying a problem situation in terms of a system, one needs to state a total objective of the system. This is much harder than it appears to be. The stated objective sometimes differs from the real objective. Reference [12] gives the example of a medical test laboratory that states their objective to perform as accurate tests as possible. Their real objective is not "accuracy" but what accuracy is good for, i.e. improving the doctor's diagnosis. If their objective is accuracy, they might sacrifice other objectives, for example spending funds wisely or containing costs. We sometimes hide our real objectives, because we believe they will not be acceptable from other's point of view. The difference between the stated objective and the real objective is that a person will not sacrifice his real objective to attain some other goal. The systems analyst should therefore identify the single goal of the system that will not be sacrificed in favor of any other goals.

The ability to measure performance goes hand in hand with stating clear objectives. We need a score to see how well the system is performing. Reference [12] uses the performance measure of a large organization as example. Should the stated goal of increasing net profit be considered as a real goal? Should the real goal not be to increase the gross profit and the growth of personnel numbers? Will the managers be willing to sacrifice a little bit of the net profit to increase the size of the organization? The true measure of performance will help us to identify the true goal of a system. One should also refer to legitimate objectives as those that have to do with the morality of the systems objectives. An objective can only be a real objective if it is acceptable from a social point of view. One cannot state objectives without a very careful examination of the consequences of these objectives.

2) The system's environment

Reference [12] defines the environment of a system as that part that is outside the system. Reference [10] defines the environment of a system as "a set of elements and their relevant properties, which elements are not part of the system but a change in any of which can produce a change in the state

of the system. Thus a system's environment consists of all variables that can affect its state. External elements that affect irrelevant properties of a system are not part of its environment. The state of a system at a moment of time is the set of relevant properties which that system has at that time". The environment determines in part how a system performs [12]. The demand for an industrial firm's product determines partly how the firm performs. Demand for the product is an example of the requirement schedule of the environment of a system.

3) The resources of the system

Resources are the means that the system uses to reach its objective. The system has control over the resources. Resources can be influenced to increase their advantage to the system. Reference [12] argues that, although a balance sheet is used to indicate a firm's resources, it does not show all the resources, for example peoples' potential. The same can be said about an income statement; it is supposed to show how the resources were used, but it does not show anything about lost opportunities. He argues that a firm needs an information system to keep track of its resources, as well as how they were used or not used in lost opportunities. Reference [12] states that "resources are the general reservoir out of which the specific actions of the system can be shaped."

4) The components of the system

Large systems need to be divided into components to aid the management scientist in determining the performance of the total system. If the performance of components can be improved, it is possible to improve the performance of the whole system. The parts or components of the system are the different activities or jobs the system has to perform. These may also be called "missions". This differs from traditionally dividing organizations into departments. Reference [12] argues that the traditional division of organizations is not a functional division of the objective of the organization. Production and Sales should be one department, since it is the production department that produces the product with which the customer is satisfied or not. The ultimate aim of component thinking is to discover those components (missions) whose measures of performances are truly related to the measure of performance of the overall system [12].

5) The management of a system

The management of a system has to deal with the generation of plans for the system. This includes the setting of the overall goals for the system, defining environment, the utilization of resources, and the division of the system into components [12].

Two authors, Ulrich [13] and [14] and Nelson [15] pay tribute the legacy of Churchman in the operations research community.

D. Ontological views of systems

This section introduces hard, soft and critical systems thinking. The first wave of systems approaches can be referred to as hard systems approaches which supported one particular human interest, namely our technical interest in predicting and controlling our environment. Second wave systems thinking involves managing debate between people so that learning may be facilitated, ideas evaluated, and plans for action developed. The third wave of systems thinking, critical systems heuristics, is concerned with subjecting assumptions in planning ethical critique [16].

1) Hard systems thinking

The term "hard systems" is used by reference [9] as an alternative to "soft systems". Hard systems thinking refers to systems engineering thinking where a systematic process of problem solving is followed. Reference [9] refers to a hard systems approach as an approach to problem solving with the assumption that the problem task is to select an efficient means of achieving a known and defined end. Systems engineers attempt to solve social problems as if they were scientific problems. Their view of a system differs greatly from the soft systems approach. True to its realistic nature, hard systems form an exact and true representation of the world. Each system can be seen as a hierarchically organized set of elements [17]. This implies that a system can be taken apart to be understood. If one is able to describe the basic elements of a system, one should also be able to describe the functionality of the system. The hard systems approach emphasizes the internal structure of the system. If the function of the system is understood, the system itself is understood.

The development of information systems has been influenced mainly by hard systems thinking. The major method of problem solving is top-down design, in which the problem is broken up into smaller, more understandable sub-problems. If the problems on the lowest level of the hierarchy can be solved, the entire problem can be solved. This approach is known as stepwise refinement [17]. Structured programming and structured design techniques are both examples of the hard system approach in information systems. The waterfall method for systems engineering views the systems development process as an objective approach that will yield objective, testable, and effective systems, answering to the problem specification. Formal problem descriptions and design methods, such as entity relational diagrams, are all part of the hard systems approach.

2) Soft systems thinking

One of the major shortcomings of the hard systems approach is that the problem is not always well defined. This makes it very difficult to reach consensus on the requirements for the new computer system. The soft systems approach views a system as a representation of the human mind to make sense of the reality [17].

Where hard systems thinking views models as representations of reality, soft systems thinking views models as aids for the development of inter-subjective understanding.

The view of human beings in a hard system environment is that of parts in a machine, or objects that could be manipulated as parts of larger systems.

The soft systems approach is holistic in that the lowest level of a system hierarchy cannot define the system. The system's purpose cannot be determined by looking at the purpose of the individual components. The systems' emergent properties give purpose to the system. In an information system environment, this means that user success, as opposed to requirements conformation, is used as a measurement of success.

Critics of the soft systems approach argue that this approach supports only one interest. It is not predicting and controlling the environment (as in hard systems thinking), but our practical interest in achieving human understanding. They argue that typical soft methodologies do not emphasize power relationships in problem situations strong enough.

3) Critical systems thinking

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, and the means in, an irreconcilable conflict and power struggle between management and workers, or system developers and users [17]. Contradictions are analyzed in detail to find prospects for alliances; different types of interventions and suggestions for change are examined and evaluated. These considerations are used to select a strategy. Actions will be performed and the situation will change, as will our conceptions and beliefs. The world, rather than people's perceptions of it, is our primary source of learning. Trade-offs in computer systems are manifestations of contradictions inherently related to the use and development of such systems.

Further discussions on the differences in these systems thinking methodologies can be found in [18]. Reference [1] give a comprehensive view of recent contribution of systems thinking to operational research and management science.

V. Empirical study: Systems thinking orientation of data warehouse practitioners

Systems knowledge provides us with many tools to better understand our information systems, but as most data warehouse practitioners do not have explicit knowledge of systems concepts, one should ask whether these practitioners use system thinking concepts intuitively, without knowing the concepts. An interpretive study was done to investigate the activities of specific data warehouse team members from a systems thinking perspective.

A. Data collection

Semi-structured interviews were conducted using 60 questions on all aspects of data warehouse development at three organizations interviewing a total of 40 participants. Questions were formulated from a systems perspective in order to better understand the intuitive systems thinking orientation of the participant. The preset questions formed the basis of each interview, although additional questions were asked to clarify answers given. Interviews were mainly conducted in Afrikaans which is the home language of the participants. All the interviews lasted between 60 and 80 minutes, and were recorded and transcribed. Since most people enjoy talking about their work, the interviews had a conversational tone. It was made very clear at the beginning of each interview that there was no correct or incorrect answer and that the researcher aimed to learn from the participants. The researcher also demonstrated some data warehousing knowledge early in the interview to establish a high standard of use of terminology. It was important to demonstrate some competence in the field in order to establish the researcher in the data warehousing paradigm. All the case studies were completed before the data was analyzed.

B. Data analysis

The nature of this study differs from typical interpretive research since typical interpretive data abstraction would not link data warehousing practices to systems thinking methodologies. A typical interpretive researcher gathers data to be able to arrive at a theory which is grounded in reality. If such an approach was followed in this research design the resulting theory would describe the practices of the data warehousing professionals in general, but not in relation to systems thinking methodologies. Therefore, the aim of the data collection was not to generate a theory by analyzing and coding the data, as is typical in interpretive methods such as grounded theory, but rather to be analyzed through pattern matching to reveal the underlying systems structures of the data warehousing practices.

Prior to data collection a mapping was done between systems thinking methodologies and data warehousing practices to guide data collection and to serve as basis for data analysis. This mapping was done on two levels of detail to guide the researcher. Firstly a high level mapping between each systems thinking methodology and data warehousing practice was done. The aim of these mappings is to give an overall perspective of a certain systems thinking methodology on data warehousing. The mappings were compiled by applying the hard, soft, and critical systems thinking on data warehousing practices.

After the high level mapping was completed a detailed level mapping was done in the form of 60 questions grouped in six categories. An example of 4 questions of this mapping is given in Table 1. The motivation for selecting these specific questions is twofold.

For each of the 60 questions a typical answer was formulated for each of the systems thinking methodologies. Since no literature on data warehousing from systems thinking perspectives could be found the researcher of this study needed to formulate these answers very carefully. The philosophical underpinning of each systems methodology was taken into account when the answers to these questions were formulated to incorporate the correct ontological assumptions

in the respective answers. Available literature on systems thinking methodologies applied to information systems in general also guided some of the formulations of the answers.

The answers to the 60 questions presented in the table need to be probable answers that can be expected from industry professionals. It also needs to be a true reflection of the specific systems methodology it represents.

In certain instances, similar answers are given for more than one methodology since the practice level of the methodologies is similar. This however does not imply that the ontological motivation for the practices is similar.

Systems thinking methodology literature does not give clear answers to many of these questions; therefore, the foundational philosophy was used to formulate an answer.

Interpretive pattern matching was used as method for data analysis. Answers to questions were carefully examined (interpreted) and compared to the pre-formulated answers in the detailed mapping table (abbreviated in Table 1). The table was used as a template to analyze the interview data in three iterations. During the first iteration an allocation was made of each answer in an applicable cell. This process was repeated for each case study. This mapping process of analysis was repeated two weeks later, without taking the first allocation into consideration. After this second allocation the two sets of tables were compared and different allocations of specific answers were investigated and corrected. A third iteration was conducted per question for all the case studies. Every question's answers were checked across all the case studies to ensure uniform allocation. All three iterations were repeated after any changes to the template were made.

An analysis report was constructed for each of the three case studies. Each report consists of two tables. The first table assigns initials to the participants, while the second table, a copy of the mapping table, was used to map the answers given by specific respondents. An example is given here as table 2 to aid the understanding of the case study reports.

In this example two people were interviewed identified as AB and CD respectively. The question asked is numbered A5 (section A, question 5) and is quoted verbatim. The response "AB: I don't know" in the hard systems thinking column means that the answer given by AB was interpreted to be a hard systems thinking answer. The specific answer is not exactly the same as the answer provided in the mapping table as a typical hard systems answer to question A5; therefore a brief summary is given of AB's specific answer namely that he does not know who owns the data warehouse. The CD in the soft systems thinking column indicates that the answer provided by participant CD is very similar to the answer in the soft systems column provided in the mapping table. The answer in the critical systems column indicates the CD said something else later in the interview, which may be mapped to critical systems thinking. An explanation of the mapping is required or his answer is too long to fit into the cell, therefore a footnote is used to present the answer given by CD. Since many of the answers were too long to quote, direct quotations are only used in crucial instances and where short answers were given.

Table 3 is an excerpt of the actual analysis table used for case study 2 with the footnotes removed.

After the data analysis, conclusions on the full table were made for each case study. For the study represented partly in Table 2 example some conclusions were:

- It is interesting to note the consensus of the answers of IM, WM and SA. It is clear that SA is more suspicious of the intentions of others than the other respondents.
- The technical warehouse team member (TS) primarily followed a hard systems approach, as specific questions on practical detail revealed.
- The team members, who leaned towards soft systems thinking, did not follow a soft systems approach throughout. Users were excluded from the data modeling process, mainly because they would not understand the models, but everyone said that star schemas are used because they are easy to understand.

When the three tables are studied from a column perspective, it is possible to detect the specific systems thinking orientation of participants and teams. Table 3 reveals a strong soft system orientation of the data warehousing team members interviewed in case study 2.

VI. Data Warehousing According to the Systems Approach

A. The data warehouse is a system

One can use the analysis tables to discuss data warehousing from a systems perspective sometimes referring specifically to the methodologies investigated. Some of the elements of such a discussion are given here:

As a soft system, the data warehouse should be defined in terms of its purpose and not its components. A data warehouse therefore, is a tool to provide management information for decision making in order to achieve the overall business objectives.

The property of providing management with accurate information which is easy to access, is viewed as an emergent property of the total system (the data warehouse). The components of the system work together to realize this purpose, rather than being the summation of the individual properties of the parts of the data warehouse.

The stated purpose of the data warehouse can be viewed from a critical systems thinking perspective by focusing the business objectives on a single critical problem area in the organization. In a typical critical systems environment, this problem area will be associated with intervention or emancipation.

This section describes the data warehouse framework presented in figure.1, according to the systems approach of [12] discussed in section III.

TABLE I
DETAILED LEVEL MAPPING OF SYSTEMS THINKING METHODOLOGIES

#	Question to data warehousing team member	Hard systems thinking	Soft systems thinking	Critical systems thinking
1	What is a data warehouse?	A data warehouse is an integrated data source to fulfil the reporting needs of business units. It consists mainly of data, metadata, and technology such as computers.	A data warehouse is a system to improve decision making in the organization. It consists of people, data and technology.	A data warehouse is a tool to affect positive change in the organization as a whole. It consists of everything required to succeed in the realization of the proposed change.
5	Who owns the data warehouse?	The development team.	More than one party, but mostly the users.	Both the involved and the affected.
9	How do you determine whether the data warehouse is successful?	Mostly a quantitative answer, or when the specification is achieved.	Qualitative answer; when the business users are satisfied.	When the problem that caused the initiation of the data warehouse project is solved.
11	Which department is responsible for the development of the data warehouse?	Typically the information technology department.	A dedicated business intelligence department.	A detailed process was followed to set up a team consisting of all the involved and affected parties. The question was asked: "Who ought to be part of the team?"

TABLE II
EXAMPLE OF DATA COLLECTED TABLE

#	Question to data warehousing team member	Hard systems thinking	Soft systems thinking	Critical systems thinking
5	Who owns the data warehouse?	AB: I don't know.	CD	CD (later)[1]

Footnote:
[1] It is really the group of people who want the change who owns the data warehouse.

TABLE III
SECTION OF ANALYSIS TABLE OF CASE STUDY 2

#	Question to data warehousing team member	Hard systems thinking	Soft systems thinking	Critical systems Thinking
9	How many users do *ad hoc* queries that they design and implement themselves? Do you encourage this?	TS: "Not many."	IM: *Ad hoc* queries are very important. SA WM	
10	Did the usage of the data warehouse influence the career paths of certain managers?		WM IM TS: "Possibly, yes."	SA

1 2	How do you know when the data warehouse is successful?	IM: When it is used. SA: When it is used. WM: When management makes better decisions. TS: When the users requests additions.	
1 3	Do you see the data warehouse a control mechanism that management uses to control how decisions are made?	IM: "No." SA: "No." WM	WM TS: It is possible.

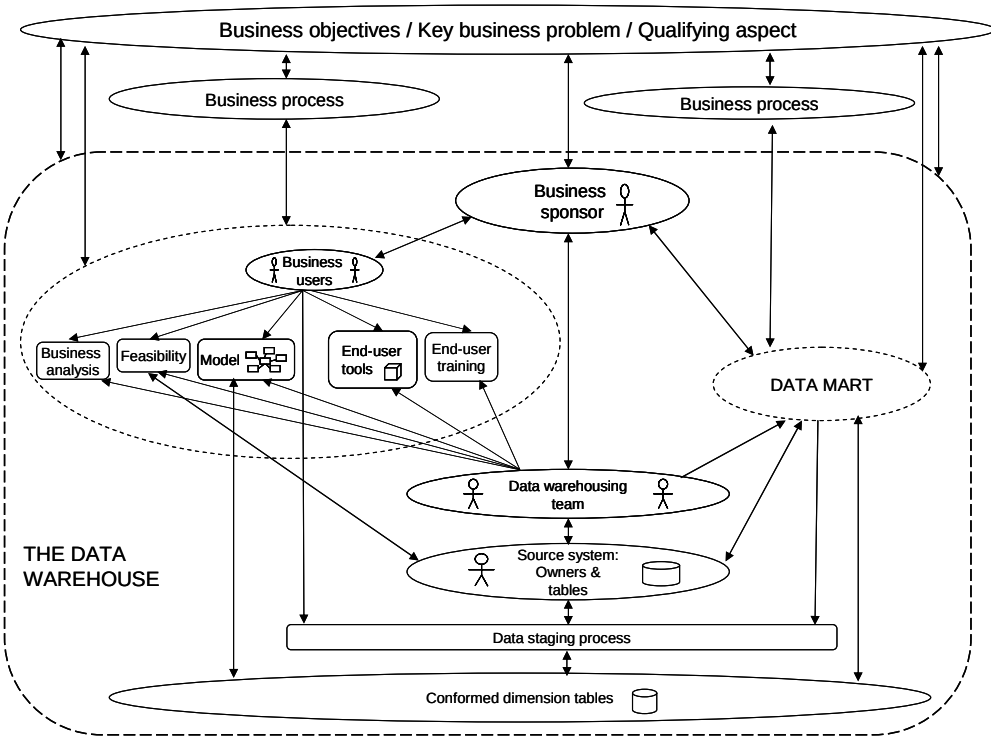

Fig. 1 Data warehouse according to the systems approach

B. The objectives of the data warehouse

The data warehouse is a subsystem of the overall system formed by the organization. When viewed through the lens of soft systems thinking: the organization's objectives (i.e. the strategic objectives of the organization), are achieved by employing the various subsystems in the organization, including the data warehouse. To ensure that the data warehouse achieves the objectives of the organization, role players from the executive committee of the organization should form part of the data warehouse. The business sponsor (described by [8]) fulfils this role. The business sponsor should be somebody who serves on the executive committee of the organization and who believes that the data warehouse can assist in achieving the overall objectives of the organization. The business sponsor should also be an influential person, able to motivate the rest of the executive committee to allocate enough resources to the data warehousing project.

The most important role of the business sponsor is to ensure that the organization's objectives are taken into account in all the activities of the data warehousing project, including the more technical activities, such as data staging. Organizations need to divide their activities into smaller areas to provide effective management. Reference [12] advocates the division of organizations into business processes, rather than traditional departments. The division of the data warehouse into data marts should follow a similar pattern. The main reason for this is the realization of overall business objectives of the organization by supporting the objectives of its subsystems. Business users from each of the business processes form an integral part of each data mart.

When these practices are viewed through the lens of critical systems thinking, one should be aware of the underlying

political agendas of the role players. The business sponsor should be chosen as somebody whose own personal objectives are compatible with the real objectives of the organization. The business sponsor should facilitate the underlying structures in the organization that should be addressed through the data warehousing project. One should also take care that the position of the business sponsor in the organizational hierarchy does not influence the effectiveness of the data warehousing project. This implies that decisions should still be reached by consensus and not be determined by the business sponsor's rank in the organization.

When a true systems approach is used, i.e. where the business objectives are truly accepted and incorporated in the data warehouse objectives, the customers of the organization will also be the customers of the data warehouse.

C. The environment of the data warehouse

Reference [12] describes the environment of a system as the factors outside the system that influence the system. These are the factors the system cannot control, but which has control over the system. In a data warehouse system, some people view the source systems as part of the data warehouse resources and others as part of the environment.

From a soft systems thinking perspective, there are four parties to consider when debating the environment of a system:

1. The organization's management and their objectives
2. The business sponsor
3. The business users
4. The source systems owners responsible for the source systems

All these parties have control over the system. The question whether the data warehouse could influence these parties varied between the case studies.

D.The resources of the data warehouse

Reference [12] states: "Resources are the general reservoir out of which the specifics of the system can be shaped." Resources are part of the system and it can be people, as well as physical instruments.

The data warehouse system has the following resources:

The business sponsor; He/she provides insight into business problems and forms the gateway to the organizational resources.

The business users; Business users are the most important resource of information about the objectives of the system.

The data warehouse team. The team includes people to assist in all components of the systems as discussed in the next section. External consultants may form part of this group. It is of the utmost importance for the data warehousing team to understand the organization's objectives.

The source systems and their owners; They are the main source of data for the data warehouse. The source system owners also provide input to quality assurance activities during the data integration process.

The conformed dimension tables; These are loaded from the source systems and form the data warehouse bus. This is similar to [8]'s view. Each data mart does not contain a copy of the data, as it is stored on a shared location. The modeling process and the interaction with the conformed dimensions are discussed in the next section.

Software tools and hardware; Software and hardware are used throughout the data warehouse to manipulate and store the data. The main data store is the conformed dimensions.

Each data mart contains fact tables with links to the dimension tables. All tables are physically stored on one or many computers

The business users are typically associated with a single business process and therefore a single data mart. The business sponsor, being a member of the executive committee, has an interest in all the data marts. The data warehousing team also has influence in all the data marts. Every data mart should not have its own data warehousing team, since it will be very difficult for such an isolated team to incorporate the overall objectives of the data warehouse and the organization in their activities. Source system input is required in all the data marts, and specific source systems may provide input to more than one data mart.

From a *soft systems thinking* perspective, the organization's objectives, supported by the data warehouse objectives, are the common factor that enables the different role players to work together. Soft systems thinking advocates an awareness of internal political aspirations and advises a method to resolve conflict and achieve consensus. The respondents of case study three stated that there is a variety of methods for reaching consensus, even in hostile situations.

The above views were supported by the respondents of case studies two and three. Most respondents agreed that people are part of a data warehouse and that the data warehouse is jointly owned by business and technical staff. Responding to user involvement, most participants agreed that senior management of business units should be involved in requirements specification. The organization reported on in case study one, draws a much tighter boundary around their data warehouse. They view a data warehouse mainly as an organized data store. They also admit that many of their problems result from this view of a data warehouse.

A *critical systems* perspective would accentuate the different agendas of the different role players. It is necessary to state for each of the role players what benefits they may gain from participating in the data warehouse. It is also important to highlight the differences in the real objectives of these role players. The technical staff's worldview, for instance, differs substantially from that of the business users (case study three). While soft systems thinking advocates the role of consensus to settle these differences, critical systems thinking admits that differences are difficult to identify and not always possible to be solved by consensus.

E. Components of the data warehouse

Reference [12] describes the components of the system as

the different activities the system has to perform. These components form the development lifecycle of the data warehouse.

Most of the participants agreed on the order of activities of the lifecycle. With reference to the data warehouse team, all but one of the participants of case studies two and three agreed on extended business user participation. They also agreed that consultants can only be beneficial when knowing and understanding the business objectives. A detailed discussion of each of the phases of the lifecycle from a systems perspective is given in [19].

F. Management of the system

The management of the system is responsible for setting the objectives of the system, defining the environment, managing the utilization of resources and dividing the system into components. In data warehouse terms, the management of the system needs to do all of the above, focused on overall quality assurance and metadata management. The performance of the system should be constantly checked to ensure that the business objectives are achieved.

The management of the data warehousing team should involve all stakeholders in the data warehouse. The group of business users shown in figure 1 is a representative group of business users. The data warehousing team comprises people trained in data warehousing and responsible for each of the components discussed in the previous section. This team may include external consultants, provided they share ownership of the organization's business objectives and therefore of the data warehouse.

The data warehousing team should include a project manager responsible for the co-ordination of all activities of the different role players. It is important to select an experienced person who is in touch with the business objectives and familiar with the key problem areas of the organization.The role players responsible for data integrity should decide on success factors and parameters for determining the success of the data warehouse. Most of the participants stated that the data warehouse is only successful if it is used to improve decision making. The project manager along with the other team members should design measures to monitor the usage of the data warehouse. The monitoring team can also identify areas where service could be improved.The project leader should ensure a high standard of technical skills relevant to the data warehouse system. This includes a detailed plan for keeping metadata updated and easily accessible. The project leader and his/her team should also ensure that the overall objectives of the organization and therefore the data warehouse are taken into account during each activity of the data warehouse system. The respondents in case study three reported that conflict can be eliminated if the responsibilities of all the role players are clearly defined.

From a *critical systems* perspective, the data warehouse will solve a specific problem and management activities will focus on aspects required to achieve this single objective. The data warehouse team will also focus on the underlying structures of the identified problem. Reference [20]'s ideas of "ought to" questions in boundary judgments would be central in the identification of role players in the data warehousing project. The critical systems thinking project leader will be aware of different agendas and motivations and will attempt to expose conflicting views to the main objective of the data warehouse.

VII. Conclusions

This paper presented data warehousing from a systems approach. It demonstrates that seemingly long forgotten ideas about system may guide us to improved user involvement and therefore greater user satisfaction and improved data warehousing success.The author hopes that some readers of this paper revisits the ideas of systems of the great systems thinkers and as academics add it to the student's learning programmes.

References

[1] J. Mingers and L. White, "A review of the recent contribution of systems thinking to operational research and management science," European Journal of Operational Research, vol. 207, pp. 1147–1161, 2010.
[2] K. Ramamurthy, A. Sen, and A.P. Sinha. "Datawarehousing, fusion and organizational effectiveness." IEEE Transactions on Systems. Man, and Cybernetics - Part A: Systems and Humans, Vol. 38, no. 4, July 2008.
[3] E. Mallach, Decision support systems and data warehousing. Boston, MA: McGraw-Hill. 2000.
[4] J. Han and M. Kamber, Data mining: Concepts and techniques. San Francisco, CA: Morgan Kaufmann. 2001.
[5] A. Sen and A.P. Sinha, "A comparison of datawarehouse methodologies." Communications of the ACM, Vol. 48, no. 3, March 2005.
[6] W.H. Inmon, Building the data warehouse. 2nd ed. New York, NY: Wiley. 1996.
[7] F.R. McFadden, J.A. Hoffer, and M.P. Prescot, Modern database management. 5th ed. Reading, MA: Addison-Westley. 1999.
[8] R. Kimball, L. Reeves, M. Ross, and W. Thornthwaite, The data warehouse lifecycle toolkit. New York, NY: Wiley. 1998.
[9] P. Checkland, Systems thinking, systems practice. Chichester: Wiley. 1981.
[10] R.L. Ackoff, Redesigning the future: A systems approach to societal problems. New York, NY: Wiley. 1974.
[11] G.M. Weinberg, An introduction to general systems thinking. New York, NY: Wiley. 1975.
[12] N.J.T.A. Kramer and J. De Smit, Systems thinking: Concepts and notions. Leiden: Martinus Nijhoff. 1977.
[13] W. Ulrich, "In memory of C. West Churchman (1913–2004) Reminiscences, retrospectives, and reflections," Journal of Organisational Transformation and Social Change, vol. 1, no. 2–3, 2004.
[14] W. Ulrich, "Obituary: C West Churchman, 1913-2004," The Journal of the Operational Research Society, vol. 55, no. 11, pp.1123-1129, Nov. 2004.
[15] H.G. Nelson, "The Legacy of C. West Churchman: A Framework for Social Systems Assessments," Systems Research and Behavioral Science, vol 20, pp. 463-473, 2003.
[16] G. Midgley, Systemic intervention: Philosophy, methodology and practice (contemporary systems thinking). New York, NY: Kluwer Academic/Plenum. 2000.
[17] B. Dahlbom and L. Mathiassen, Computers in context: The philosophy and practice of systems design. Oxford: NCC Blackwell. 1993.
[18] M.C. Jakcson, Systems methodology for the management sciences. New York, NY: Plenum. 1991.
[19] R. Goede, A framework for the explicit use of specific systems thinking methodologies in data-driven decision support system development. Pretoria: UP. 2004.
[20] W. Ulrich, "Critical heuristics of social systems design," in Critical systems thinking: Directed readings, R.L. Flood and M.C. Jackson, Eds. Chichester: Wiley. 1987. pp.103-116.