AGILE BUSINESS INTELLIGENCE DEVELOPMENT CORE PRACTICES

Master’s (two year) thesis in Informatics (30 credits)

Surendra Devarapalli

Summer 2013: 2013MASI01
Title: AGILE BUSINESS INTELLIGENCE DEVELOPMENT CORE PRACTICES

Publishing Year: 2013

Author: Surendra Devarapalli

Supervisor: Rikard Lindgren

Abstract

Today we are in an age of Information. The systems that effectively use the vast amount of data available all over the world and provide meaningful insight (i.e. BI systems) for the people who need it are of critical importance. The development of such systems has always been a challenge as the development is outweighed by change. The methodologies that are devised for coping with the constant change during the system development are agile methodologies. So practitioners and researchers are showing keen interest to use agile strategies for the BI projects development.

The research aims to find out how well the agile strategies suit for the development of BI projects. The research considers a case study in a very big organization as BI is organization centric. There by assessing the empirical results that are collected from interviews the author is trying to generalize the results. The results for the research will give an insight of the best practices that can be considered while considering agile strategies and also the practical problems that we may encounter on the journey. The findings have implications for both business and technical managers who want to consider agile strategies for the BI/DW development projects.

Keywords:

BI, Agile strategies, Scrum, Extreme programming, Data warehousing, Analytics
# Table of Contents

1 INTRODUCTION ........................................................................................................... - 1 -
  1.1 BACKGROUND ........................................................................................................ - 1 -
    1.1.1 Relation to Informatics ...................................................................................... - 2 -
  1.2 STATEMENT OF PROBLEM ................................................................................... - 2 -
  1.3 PURPOSE OF THE STUDY ..................................................................................... - 2 -
  1.4 RESEARCH QUESTIONS .......................................................................................... - 3 -
  1.5 TARGET GROUP ..................................................................................................... - 3 -
  1.6 DELIMITATIONS ..................................................................................................... - 3 -
  1.7 EXPECTED RESULT ............................................................................................... - 4 -
  1.8 THE AUTHOR’S EXPERIENCE AND BACKGROUND ............................................. - 4 -
  1.9 STRUCTURE OF THE THESIS ............................................................................... - 4 -

2 RESEARCH DESIGN ...................................................................................................... - 7 -
  2.1 RESEARCH PERSPECTIVE ..................................................................................... - 7 -
    2.1.1 Motivation for choosing hermeneutics: .......................................................... - 8 -
    2.1.2 Arguments for qualitative method: ................................................................. - 8 -
  2.2 RESEARCH STRATEGY ............................................................................................ - 9 -
    2.2.1 Research Approach ....................................................................................... - 10 -
  2.3 RESEARCH SETTING .............................................................................................. - 11 -
  2.4 DATA COLLECTION PROCEDURES ........................................................................ - 11 -
  2.5 DATA ANALYSIS PROCEDURES .......................................................................... - 12 -
  2.6 STRATEGIES FOR VALIDATING FINDINGS ........................................................... - 13 -
  2.7 RESULT PRESENTATION METHODS AND REFERENCING METHODS .................. - 13 -

3 THEORETICAL STUDY .................................................................................................. - 14 -
  3.1 KEY CONCEPTS ..................................................................................................... - 14 -
  3.2 SUBJECT AREAS RELEVANT FOR THE RESEARCH ........................................... - 14 -
  3.3 PREVIOUS RESEARCH .......................................................................................... - 15 -
  3.4 RELEVANT LITERATURE SOURCES ..................................................................... - 16 -
  3.5 BI DEVELOPMENT .................................................................................................. - 18 -
    3.5.1 BI .................................................................................................................... - 18 -
    3.5.2 Case for Agile BI development ....................................................................... - 25 -
  3.6 AGILE SYSTEM DEVELOPMENT .......................................................................... - 26 -
    3.6.1 Need for Agile Methods .................................................................................. - 26 -
    3.6.2 Agile Methodologies ...................................................................................... - 28 -
  3.7 AGILE STRATEGIES ............................................................................................... - 32 -
  3.8 SUMMARY OF THEORETICAL FINDINGS ............................................................. - 39 -
    3.8.1 Typical transactional processing systems development vs. the BI assignment development ................................................................. - 39 -
    3.8.2 How do we implement agile strategies in BI development? ......................... - 40 -
  3.9 ARGUMENTS FOR AN EMPIRICAL STUDY ........................................................... - 45 -

4 EMPIRICAL FINDINGS .................................................................................................. - 46 -
  4.1 PURPOSE ................................................................................................................. - 46 -
  4.2 SAMPLING ............................................................................................................. - 46 -
    4.2.1 The Respondents ............................................................................................ - 46 -
  4.3 THE INTERVIEWS .................................................................................................. - 46 -
  4.4 INFORMATION FROM THE VARIOUS RESPONDENT’S ...................................... - 48 -
  4.5 EMPIRICAL RESEARCH RESULTS ....................................................................... - 52 -

5 ANALYSIS AND RESULT ............................................................................................... - 55 -
  5.1 ANALYSIS .............................................................................................................. - 55 -
  5.2 RESULT SUMMARY ............................................................................................... - 61 -

6 DISCUSSION ................................................................................................................. - 65 -
  6.1 CONCLUSIONS ....................................................................................................... - 65 -
  6.2 IMPLICATIONS FOR INFORMATICS ..................................................................... - 66 -
List of Figures:

Figure 1 Pictorial representation of BI Terms ................................................................. - 2 -
Figure 2 Thesis Structure ................................................................................................... - 6 -
Figure 3 Subject areas and their relevance to the research question .................................. - 15 -
Figure 4 BI Framework “Getting data in & getting data out”(Watson & Wixom, 2007)......... - 21 -
Figure 5 A typical DW/BI Architecture ............................................................................. - 24 -
Figure 6 Agile/Lean way of working Agile DW/BI .............................................................. - 35 -
Figure 7 Agile/Lean way of working .................................................................................. - 44 -
Figure 8 Core practices for bringing agility to BI development projects ............................. - 64 -

List of Tables:

Table 1 Comparison of Qualitative & Quantitative methods .............................................. - 10 -
Table 2 Sample search phases ........................................................................................... - 16 -
Table 3 Transactional systems vs. Data warehouse systems Characteristics ...................... - 19 -
Table 4 BI Applications vs. Standalone applications ......................................................... - 23 -
Table 5 BI development stages (Ion C. Lungu, 2005) ......................................................... - 24 -
Table 6 The main characteristics of each one of these eight ASDMs (Goede & Huisman, 2010) ...... - 32 -
Table 7 Various practices for BI/DW development .............................................................. - 37 -
Table 8 Application of ASDM principles in DW adapted (Goede & Huisman, 2010) ............ - 38 -
Table 9 Empirical Findings vs. Theoretical results .............................................................. - 60 -

Abbreviations Used

<table>
<thead>
<tr>
<th>BI</th>
<th>Business Intelligence</th>
</tr>
</thead>
</table>
Thanks to ALMIGHTY

The opportunity to work in a tremendous organization is invaluable and been a wonderful experience. I would like to thank IT Company for giving me this exciting opportunity.

I am very much thankful to everyone who has supported me in writing this thesis. First of all I want to thank my academic supervisor Rikard Lindgren and supervisor at IT Company, Steve Van Hoyweghen for their valuable support in completing this thesis. I want to thank especially the Line Manager, Bogdan Bunea for encouraging me to finish my thesis at an earlier phase and for providing me with valuable contacts. I am greatly indebted to my Line manager, Christina Tröfast and my whole BI team who are always keen to help me with necessary information for the thesis.

Finally I would like to thank my family who has supported me all through the journey.

Regards
Surendra Devarapalli
1 INTRODUCTION

This chapter presents the background of the research area and the research questions that this thesis is trying to acknowledge along with the significance of the research. It presents the tentative structure of the thesis. The chapter gives the relation of the research to informatics along by outlining the delimitations.

1.1 Background

With the advent of new technologies in the fields of information and communication there is a huge proliferation of data. The process of incorporating such huge amount of data into the business information systems for aiding the management decisions has become a hefty task in the last decade (Cukier, 2010; Jim Highsmith, 2002; Wilson, 2012). If this substantial data is used selectively and send to specific people, the organizations can build competencies for them (RODRIGUES, 2002). There is a high need for the managers to react accordingly to the rapid changing world, globalizations and gigantic organizations and make decisions(Sullivan, 2002). For this purpose BI systems can be used in strategic and managerial processes (RODRIGUES, 2002).

BI is a set of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information used to enable more effective strategic, tactical, and operational insights and decision-making. A narrow definition is used when referring to just the top layers of the BI architectural stack such as reporting, analytics and dashboards (Evelson, 2010).

Enhancements of the BI systems have been a huge and daunting task for the most IT organizations. In a recent survey by Forrester Research claimed that BI decision-makers realized that 70 percent faced business requirements change monthly or even more frequently. The fact is that the traditional BI processes have never been considered as agile or responsive or flexible (Caruso, 2011).

BI development is often incremental in nature. Unlike transactional processing systems like OLTP there is always a need for the BI systems to change for improvements. As the BI systems give strategic support for taking management decisions they should continuously change in accordance with the organizations internal requirements. For this BI development project there need to be very good plan before starting the BI project (Ko & Abdullaev, 2007).

For to make a BI application to embrace change requirements in an ever-changing world the application of the agile processes and management models like Scrum or Extreme programming are being discussed.(Knabke & Olbrich, 2011; Schwaber, 1997).

This research focuses on BI development and is primarily involved in two main activities that are getting data in and extracting data out. Getting data is traditionally referred to as Data Warehousing that is gathering data from a set of source systems to an integrated data warehouse and aligns them accordingly with the requirement.
The unit of analysis or the phenomenon that this thesis is going to speak about is business intelligence development.

Figure 1 Pictorial representation of BI Terms

1.1.1 Relation to Informatics

Informatics studies the interaction of information with individuals and organizations, as well as the fundamentals of computation and computability, hardware and software technologies used to store, process and communicate digitized information. It includes the study of communication as a process that links people together, to affect the behavior of individuals and organizations (Michael.F, 2002). The main focus of Informatics is the transformation of Information either by Computation or Communication whether by Organisms or Artifact’s (University of Edinburgh, 2010). This thesis is on BI whose very idea is about transformation of Data to Information and in turn Information to Knowledge. This forms as the convergence point for Informatics and this thesis.

1.2 Statement of Problem

This thesis investigation starts on the Title “Agile business intelligence development Core Practices”. BI development is Incremental in nature with typical short cycles. As such it seems that agile methods suits extremely well for the development process of BI.

Agile is organization specific so every organization has its own way of describing and implementing agile methods for their BI development projects. Hence there is a list of practices recommended for making the development process agile. With this in context there is limited to no information on how to adopt agile methods for Specific BI development projects.

1.3 Purpose of the study

The purpose of the study is to study the agile strategies and show how well they suits for BI development. This thesis is considering one of the biggest IT organizations in Sweden as a context for applying agile strategies to BI development. The study intends to reduce the gap between the current way of working and proposed way of working. The purpose of the study at the end is to suggest the BI community about how agile strategies can be adopted for BI development.
1.4 Research Questions

Research question:

Q. How can we adopt the agile way of working for BI development?

To make the process for answering the main research question it is divided it into sub questions. The research will try to answer sub questions so that it leads to the end that the main research question is answered.

Sub Questions:

1. What are the Similarities and differences between the Traditional transaction processing systems development and BI systems development?
2. How to implement agile strategies in BI development?

The 1st sub question gives the picture on how the traditional transactional processing systems development and BI systems development match to each other. By this it’s possible to see if the same agile methods used for transactional processing systems development can be used for BI systems by implementing necessary changes. The 2nd sub question explains how to implement pragmatically agile strategies in BI development using some core practices.

1.5 Target Group

Researchers

The dissertation results will help the researchers who are working in the field of agile strategies as well as the researchers who are working in the field of BI development. By this, the researchers can start working on the risks that are identified while implementing agile strategies for BI development. This enables the investors in the BI projects to be more profitable with every short deliverables of working software.

Managers in IT organizations

This thesis also helps the Managers and Business Executives who want to take decisions on developing BI projects. This research results illustrates the risks and benefits of both approaches namely traditional way of development and agile way of development. Further the managers can decide which BI project to be developed and in what approach.

1.6 Delimitations

This research mainly considered IT Companies as the reference. The author discusses about agile strategies that are used in specific to that field. Here he suggests the agile way of working for BI at IT organization. This does not end up making suggestions for replacing current development of BI with Agile strategies for other organizations as a whole because agile is organization specific so it differs from one other.
1.7 Expected Result

The research presents the similarities and differences between the classical developments of transactional systems using object oriented programming languages and BI development Life cycle.

The expected outcome from this study involves the reflections based on experiences which are acquired from a real time BI project and on an existing literature study which shows How to adopt agile methods for BI development. These reflections can serve to suggest the ways to bridge the gap between current way of working and the proposed agile way of working. This may involve in Suggestion of tools, techniques and methods adaptations if required.

1.8 The Author’s Experience and Background

The author has very good theoretical foundations with the courses on BI and Agile methodologies which are completed during his Master’s and Bachelor’s Studies.

The possibility of working in a significant organization for the thesis enabled him to get in contact with real-time people involved in the area of study. This also helped in facilitating the easy reach of the personnel for attaining the qualitative data.

1.9 Structure of the Thesis

The research involves a series of steps to arrive at the results and conclusion. The thesis work is organized into 6 different chapters. They are

The First chapter is introduction which gives the background of the research area and the arrival of the research questions with connection to the informatics. It also deals with significance of the research and helps to split the main research question into sub research questions.

The Second Chapter is research design which includes the research strategy, data collection methods and procedures. It also includes the data analysis procedures and the strategies that are used to analyze the validity of the findings.

The Third Chapter is theoretical study which forms a strong foundation for the research that is to be done. Relevant literature sources are found, the key words are identified and exhaustive study is made in order to show up the theoretical results which form as a basis for the empirical survey that is to be carried after.

The Fourth Chapter is empirical findings which emphasize the need for empirical findings. The data is collected using empirical methods like interviews and questionnaires.

The Fifth Chapter is analysis and results which starts with analysis of the theoretical and empirical results and ends with presenting the results.
The sixth Chapter is discussion and conclusions which includes the discussion on the methods and results that leads to the conclusion. This chapter ends with suggesting future research scope.

At the end we can find the References section (i.e. Harvard referencing Style).
Figure 2 Thesis Structure
2 RESEARCH DESIGN

The research design provides the framework for the collection and analysis of data. This explains the criteria that are employed for conducting the business research. This chapter concludes by providing the strategies for validating the findings.

2.1 Research Perspective

The knowledge that is created after the research can be of different kinds. They can be either normative or descriptive. By defining the knowledge the researcher is going to create, it sometimes helps us to justify their actions (Gilje & Grimen, 1993)

There are mainly two scientific perspectives namely positivism and hermeneutics for conducting research. Natural science dominantly has positivist view of understanding and the humanities have the hermeneutics way of understanding. Hermeneutics aims at understanding and explaining meaningful concepts. As this research is aimed in creating comprehensive knowledge through interpretation of text the author feels that hermeneutics is the most relevant approach as it involves interpretation of text gathered from company and interviews (Verbeek, 2003)

There are 3 types of Epistemological position and they are positivism, realist and interpretive. An epistemological issue concerns with the question of what should be considered or regarded as the acceptable knowledge in a discipline. An important question that we need to consider is that whether the social world can and should be studied according to the same principles, procedures, and ethos as the natural sciences. The position that affirms the importance of imitating the natural science is constantly associated with the position known as positivism (Bryman, 2012)

Positivistic perspective:

Positivism is an epistemological position that advocates the application of the methods of natural science to study the social reality and beyond this principle. But the term is stretched beyond this principle, and also it differs between various others. But positivism includes the following principles.

1. Principles of phenomenalism: The knowledge that is perceived by the senses can only be considered to be as the warranted knowledge.
2. The purpose of the theory is to generate hypotheses that can be tested and that will thereby allow explanations of laws to be assessed (the principle of deductivism).
3. Knowledge is arrived at through the gathering of facts that provide the basis for laws (the principle of inductivism).
4. Science must be conducted in a way that is value free (i.e. objective).
5. There is a clear distinction between scientific statements and normative statements and a belief that the former are true domain of the scientist (Bryman, 2012)

Hermeneutic perspective:

Hermeneutics is the theory of understanding. For a person to get a meaningful understanding of the information one should start with the fundamental theory of meaning and understanding and interpretation available i.e. with hermeneutics. (Introna) The social enquiry is characterized by the “double hermeneutic”. The post-positivist views gained epistemological ground by acknowledging the first half of the structure of inquiry, namely that a science theory and findings are shaped by the investigators interpretive framework of assumptions, conventions and purposes. The second half of this double hermeneutics is that the characteristics of the human action and emotions are structured by the social reality (Richardson & Woolfolk, 1994) With the help of the hermeneutics the researchers can gain knowledge by interpreting texts, language and design. The researcher learns and develops answers by collective experiences.

2.1.1 Motivation for choosing hermeneutics:

The research is considering the hermeneutic perspective for this research. This is for the fact that hermeneutics explain meaningful concepts by interpreting text, language and artifacts. As hermeneutic approach emphasizes on the location of the interpretation within the specific social and historical context, would seem to logically conclude that the analysis of texts is in conversant with the context. Some authors identified an approach for the interpretation of the company documents that which they described as ‘Critical hermeneutic Approach’ (Phillips & Brown, 1993). In this research for getting insight of the methods and strategies used in the organization in this context namely IT company in order to suggest pros and cons of implanting them in a new way the author uses hermeneutics. Hermeneutics is used here in this research for interpreting company documents and text that is gathered by interviews. With all this in context the author feels hermeneutics is the most relevant approach.

2.1.2 Arguments for qualitative method:

There are generally two methods for conducting research they are quantitative methods and qualitative methods (Yin, 1994). In this research as the author has confined to the hermeneutic perspective for the creation of knowledge, qualitative method is the most suitable method. Every researcher chooses the method of research by taking into account the type of knowledge he is going to create. Both methods are equally good but it’s only a matter of context for which they are applied. The research involves interpreting the company documents and the texts that is acquired from the interviews. Hence as the data created is qualitative in nature qualitative method is used.
2.2 Research Strategy

Quantitative and qualitative can be taken as two distinctive clusters of research strategies. A research strategy means that the general orientation for the conduct of business research.

**Quantitative:** Quantitative strategy can be understood as that research strategy which emphasizes quantification in the collection and analysis of the data.

**Qualitative:** Qualitative strategy can be understood as the research strategy that emphasizes the words rather than the quantification in the collection and analysis of the data.

There are several research designs for conducting research. Bringing the research strategy and the research design together plays an important role in pursuing the research. The author is outlining five of the research designs. They are Experimental design; cross-sectional design or social survey design; longitudinal design; case study design; and comparative design. According to (Yin, 1994) experimental design best fits only for quantitative study as it depends on values and calculations.

With the research being involved with the practices that are being implemented in an organization for business the more suitable research design here is the case study design. This research is considering case study design. The case study involves in intensive analysis of a single case. A case studies complexity depends upon the nature of the case in question (Stake, 1995) The case study design approach is very popular and widely used research approach. A case can be a single organization or single location or person or a single event.

The case study approach is different from the other approaches because the focus is on a bounded situation or system, an entity with a purpose and functioning parts. The main emphasis is on the intensive examination of the setting. For detailed examination of a case it is very evident that the methods like participant observation and unstructured interviewing are helpful. From the findings of the case study we can gain insight into that case (Bryman, 2012)

Case study approach is considered to be a very robust research method when a holistic and in-depth investigation of it is required. Case studies observe the data at micro level. The researcher can choose between the single case and multiple case designs depending on the issue in the research question. If there are no cases of replicability then it is better to use single case design. Multiple case designs can be implemented when there are numerous cases of replicability (Zainal, 2007) According to (Yin, 1994) the generalization of the results either from the single case or multiple case design depends more on theory rather than population.

As agile methods and BI development methods are considered from a specific IT company the case study approach is best suited.
Category of Case Study:

There are various types of case studies. Yin (1994) presents three categories namely exploratory, descriptive and explanatory case studies. In exploratory studies set to study the phenomenon in the data which seems interesting to the researcher. It may start with very general questions and that helps for the studying of the phenomenon. In this case study a small field work and small scale data collection may have been conducted prior to the research questions or hypothesis is made for the research.

Descriptive studies are those that study the phenomenon that occurs while studying the data. The goal set by the researcher is to describe the data that occur. The exploratory studies are that which study the data at surface level and deep level in-order to explain the phenomenon in the data (Zainal, 2007) In this research the author is using the descriptive studies as the research question that is How to adopt Agile way of working for BI development. The author feels that descriptive studies best suit for above research.

According to the case study approach the Phenomenon that the research is going to say something is the BI development.

2.2.1 Research Approach

It is always important to define the nature of research that the author is going to undertake. The role of theoretical and empirical study has to be clearly defined. This can be found by the approach we are going to consider that is if it is the deductive approach or the inductive approach. Deductive research constructs the hypothesis based on the theory and examines it using the empirical study (i.e. testing of theory). An Inductive research infers the empirical findings for revising the theory in a particular domain (Bryman, 2012)

In this Research the author is using the Inductive approach. That is the empirical data obtained by using the qualitative methods like interviews and questionnaires are used. The results from the empirical study and the findings from the theoretical study are compared to each other. The author is using Inductive approach for the research.

<table>
<thead>
<tr>
<th></th>
<th>Quantitative</th>
<th>Qualitative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Principal orientation</td>
<td>Deductive; Testing of theory</td>
<td>Inductive; Generation of theory</td>
</tr>
<tr>
<td>Epistemological orientation</td>
<td>Natural science model, in particular positivism Objectivism</td>
<td>Interpretivism Constructionism</td>
</tr>
</tbody>
</table>

Table 1 Comparison of Qualitative & Quantitative methods
2.3 Research Setting

The research setting considered for this study is at a large multinational IT company which has nearly 5000 employees working across the globe. The author tried to reflect on the experiences gained from the BI projects conducted within that BI team and as well as from the knowledge gained from the theoretical study. This allowed to perceive the agile way of working from two different viewpoints one from theory and the other from practical application as well. This allowed juxtaposing both the ways of working which helped to suggest some efficient ways of working in agile way for BI development.

2.4 Data Collection Procedures

According to (Stake, 1995) the primary methods for qualitative data collection are interviews, data documents and observations and analyzing the documents. The researcher can start by picking some data randomly and start reading so that he gets some understanding of the case. The data collection strategy is determined by the question of the study which we answer at the end of the research and by determining the sources of the data which yield best data for answering the question. Researchers are encouraged to collect data using more than one method of data collection which enhances to test the validity of the findings.

**Interviews:** Interviews range from highly structured with specific questions and the order is determined ahead of time. The unstructured interviews where the questions and the order are not determined but they have the subject areas in their minds to explore. Most of the interviews fall in between these two. They are called as semi-structured interviews in which there are some specific questions and structure for the interview (Merriam, 2002)

**Observations:** One of the major means of collecting data is through observations. The data obtained by observation gives the first hand impression of the data rather than second hand data obtained by the interview. There are two types of observer’s namely complete observer and the active participant. Complete observer is only observing the case but the observed group is not aware of this observer. But the active participant is a part of the observation group and he observes while he actively participates in the organization. Observation is a very good technique when the researcher desires to have a firsthand knowledge (Merriam, 2002).

**Documents:** The third major source of the data is documents. The entire study can be made around documents. These documents can be any type either oral, text, visual or cultural artifacts. Public records, office documents personal material are the different kinds of documents available for the researcher (Merriam, 2002)

Interviews, Observations and the documents are the three traditional sources of the documents. But with the advent of the internet and telecommunication technologies we can collect data from variety of sources through them.

The author has been the using all the above mentioned techniques for collection of the data like Interviews, Observations, Documents and internet etc. Data Collection for research can be of two types. They are primary data and secondary data. The primary
data are collected by using the traditional methods like interviews, questionnaires, and observations (J.W. Creswell, 2007) For some researchers it is possible to collect the data collected by other researchers this data is called as secondary data. The secondary data can be of any type like official records, administrative statistics or the records that the organizations maintain to keep track of the organizational changes etc. (Hox & Boeije, 2005).

While going for the data collection methods the author tried to answer few questions for getting insight into the methods to be chosen for collecting data. They are as follows

1. What are the resources and accessibility constraints I have for my research?
2. What information is already available and what need to found?
3. What are the appropriate methods for this kind of research in specific?

The author has taken the help of the secondary data for gaining the basic understanding of the research area. Then he depends on the primary data which is collected by interviews, observation etc. With the secondary data available the author formed a strong theoretical base and then by effectively using the primary data collected by interviews and observations he conducted to validate the theoretical findings.

2.5 Data Analysis Procedures

Data collection and analysis procedures are developed in an iterative procedure in a case study (i.e. in contrast with the experimental and surveys). This is very much helpful because the theory developed is very well grounded with the empirical evidence (Hartley, 2004)

There is no particular moment of data analysis. Analysis of the data is a matter of giving meaning to the first impressions and as well as final compilations (Stake, 1995) There are several levels of data analysis in the case study which is qualitative. One of the useful actions that will help us a lot is that to arrange the data in a chronological order. This helps us to understand the data and present it in a descriptive manner. It helps to create or develop theories, models and draw inferences (Merriam, 1998)

Primary data:

The primary data is collected using the data collection techniques like questionnaires and interviews. The responses are documented from the questionnaires and a note is taken while taking the interview. Some interviews are conducted using office communicator and some face to face interviews. The data analysis method that the author will be using is comparative in nature (Stake, 1995)
Secondary Data:

The analysis of secondary data and theory is done by the following method. Most of the secondary data is found using the University of Borås Summon tool and Google scholar. Initial selection of articles is done by reading the abstract. Then interesting and relevant articles are shortlisted and the details are studied. By skimming the data in the articles several articles are filtered and around 25 articles are shortlisted and used for the final analysis.

2.6 Strategies for Validating Findings

For analyzing the case study approach that how well it fits in the research design we have criteria like Measurement validity, internal validity, external validity, ecological validity, reliability and replicability. It also depends on how well the researcher finds the above criteria relevant. Researchers like (Yin, 1994) consider that these criteria are very relevant and try to enhance their ways in order to enhance their ability to meet the criteria.

The criteria like reliability, validity and generalizability are the different kinds of measurements of quality, rigor and wider potential of research. Validity refers to whether you are observing, identifying, or “measuring” what you say you are (Bryman, 2012)

The author used triangulation method, External Reliability, Internal Reliability, and Internal Validity for validating findings. The reasons for using them are the following.

“Triangulation is a validation procedure where researchers search for convergence among multiple and different sources of information to form themes or categories in a study” (John W. Creswell & Miller, 2000)

External Reliability: This describes the degree to which the study can be replicated.

Internal Reliability: This means that there is a consensus among the observer’s (which are more than one) about what they hear and see.

Internal Validity: This means that to check if there is a good match between researcher's observations and the theoretical ideas develop (Bryman, 2012)

2.7 Result Presentation methods and referencing methods.

The results are presented more in text format. The results of the case study have been reported with much care. The results are carefully documented using text and diagrams and tables. The author used Harvard system of referencing. This involves in citing with the name of the author and along with the year.
3 THEORETICAL STUDY

3.1 Key Concepts

BI is a set of methodologies, processes, architectures and technologies that transform raw data into meaningful and useful information used to enable more effective strategic, tactical, and operational insights and decision-making. A narrow definition is used when referring to just the top layers of the BI architectural stack such as reporting, analytics and dashboards (Evelson, 2010).

BI is a broad category of application programs and technologies for gathering, storing, analyzing and providing access to data to help enterprise users make better business decisions. BI applications include the activities of decision support, query and reporting, online analytical processing (OLAP), statistical analysis, forecasting, and data mining (Rossetti, November 2006).

BI is the process of gathering information in the field of business. It can be described as the process of enhancing data into information and then into knowledge. A popularized umbrella term used to describe a set of concepts and methods to improve business decision making by using fact-based support systems. The term is sometimes used interchangeably with briefing books and executive information systems (Mirum.net, ND).

Data Warehousing: Data warehousing is a collection of decision support technologies, aimed at enabling the knowledge worker (executive, manager, analyst) to make better and faster decisions (Chaudhuri & Dayal, 1997).

“A software development method is said to be an agile software development method when a method is people focused, communications-oriented, flexible (ready to adapt to expected or unexpected change at any time), speedy (encourages rapid and iterative development of the product in small releases), lean (focuses on shortening timeframe and cost and on improved quality), responsive (reacts appropriately to expected and unexpected changes), and learning (focuses on improvement during and after product development)” (Qumer & Henderson-Sellers, 2008).

Agile methodologies, such as extreme Programming (XP), have been touted as the programming methodologies of choice for the high-speed, volatile world of Internet and Web software development. Although creators of agile methodologies usually espouse them as disciplined processes, some have used them to argue against rigorous software process improvement models such as the Capability Maturity Model (CMM) for Software (SW-CMM)(Paulk, 2002).

3.2 Subject areas relevant for the research

This section deals with the introduction of the subject areas and their relevance to the research questions.

The following are observed to be the relevant subject areas for answering the sub questions and in-turn the main research question. They are
3.3 Previous Research

There is a considerable amount of specific research done aiming at the “Development of BI Using Agile Strategies”. The search for the literature in this Agile BI results mostly in Agile Data warehousing and considerable results for Agile Analytics. There are not many specific development methodologies in the market but only a collection of best practices for data warehousing and data base developments.

Development of BI is always incremental in nature unlike building of regular transactional processing systems (OLTP). Building a BI system actually never ends as the system always need to be revised for adding new requirements or functionalities (Ko & Abdullaev, 2007)

Agile development methodologies have been very significant in the recent years. But there has been no specific research done on how to use these methodologies for Specific BI developments (Cao, Mohan, Xu, & Ramesh, 2009). The agile methods very well suits for the development of systems that can cope up with very volatile environments and changing requirements. Most of the research has been focused on
proposing to develop the Data warehousing architecture in agile way. But very little contribution has been done on how to develop the process of creating and changing the whole BI system (Knabke & Olbrich, 2011).

Ken collier is one of the first to introduce agile methods for the development of Data warehousing, business Intelligence and agile analytics. In his book agile Analytics in 2011 has successfully adopted agile techniques for Data warehousing and BI to create agile analytical style (Collier, 2011).

Developing of BI using agile strategies does not have clear methodology to adopt in the research setting under consideration for implementation. As the BI development is organization specific and centric there is a need for to propose a methodology for developing BI using agile strategies in that context.

3.4 Relevant Literature sources

<table>
<thead>
<tr>
<th>Some keywords and Phrases used for searching</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI Development</td>
</tr>
<tr>
<td>Agile Strategies</td>
</tr>
</tbody>
</table>

Table 2 Sample search phases

The author has taken the aid of the book “Agile software development ecosystems” written by High smith for acquiring the basic knowledge about the agile principles and the dominant agile methodologies that are currently in use (Jim Highsmith, 2002) This outlines seven agile approaches that are commonly followed like Scrum, Dynamic System Development Method (DSDM), Crystal Methods, Feature-Driven Development (FDD), Lean Development (LD), Extreme Programming (XP), and Adaptive Software Development (ASD).

Papers that have topics related to the Lean Methodology are considered. The book on “Lean Integration” by John Schmidt helped the author to understand the Lean Integration System. It gives the insight about the Lean Methodology from the inception to the changes and it current way of operating in the industry (Schmidt & Lyle, 2010)

The above keywords that are listed in the above table are not the exhaustive list, but the collection of the important keywords that are used in the search.

BI Development:

The article “BI: An Analysis of the Literature” by Zack Jourdan, R. Kelly Rainer, and Thomas E. Marshall, 2008 provided with an exhaustive coverage of the literature
published from 1997 to 2006. The Journal article by Watson, 2007 on “The Current State of BI” helped the author for having an understanding of the BI and the current state of BI. The article by Rina Fitriana, Eriyatno, Taufik Dijatna on “Progress in BI System research: A literature Review” helped the author to have a very good literature on the BI system research that was conducted from 2000 to 2011.

The Basic developmental stages of the BI development and their characteristics were initially drawn from the research article “BIDM: The BI Development Model” by Catalina Sacu, Marco Spruit, 2010. The relevant literature source “A framework for accessing an enterprise business intelligence maturity model (EBI2M).: Delphi study approach” by Min-Hooi Chuah and Kee-Luen Wong, 2012 outlined nearly 9 BI development models. This contrasts the major differences and their issues of focus among them. The article “The Dynamic structure of management support systems: Theory Development, Research focus, and Direction” provided with architectural model with the constructs that are observed to be impacting the success of the decision support systems.

Agile Strategies:

Lean Integration by John G. Schmidt gave the complete knowledge of Lean integration and lean software development and their background and current status. Lean software development is an agile approach that translates Lean manufacturing principles and practices for the software development domain (Schmidt & Lyle, 2010).

Scrum is not a methodology it is a framework. The paper by Henrik kniberg on Scrum and XP from trenches gave an understanding of scrum and how to integrate XP practices and scrum and the how it has been practically implemented in their organization. The Scrum Guide by Jeff Sutherland and Ken schwaber is considered as complete reference for the knowledge about scrum. The book by Kent Beck on Extreme programming titled “Extreme programming explained” is used a reference guide for the Extreme programming.

Data warehousing

Data warehouse is simply a single, complete and consistent store of data obtained from a variety of sources and made available to end users in a way they can understand and use in a business context (Devlin, 1996). The classical view of development is to have all the requirements before designing and developing an application. But for data warehousing systems it’s not always possible to have all the requirements enunciated before the development of the information systems. Hence the data warehouses are developed iteratively. The classical operational environment is developed using the SDLC- which requires the requirements gathering at first followed by analysis and design and then programming followed by testing and implementation. That is for when we know all the requirements details but the development of a data warehouse is done in an entirely different way (Inmon, 2005).

Agile Analytics

Agile Analytics is a user-value–driven approach in which high-valued BI capabilities drive the evolutionary development of the data warehouse components needed to
support those capabilities (Collier, 2011) The book *Agile Analytics* by Ken collier is considered to be a very good literature asset for this research on developing BI by agile methods.

The author has found relevant literature sources from the various journals, books and conference papers. Online databases were searched for in the University of Boras and Google scholar and internet.

*Databases*: EBSCO-Business Premier, ACM Digital Library, Blackwell Synergy, SAGE Journals, ScienceDirect, SpringerLink, and Wiley Online Library

*Journals*:
- Communications of the ACM
- Communications of the Associations for Information Systems
- European Journal of Information Systems
- Information Systems Journal
- Information Systems Research
- Journal of Information Technology
- Journal of Management Information Systems
- Journal of Strategic Information Systems
- Journal of the Association for Information Systems
- MIS Quarterly

### 3.5 BI Development

#### 3.5.1 BI

A transactional processing system is a computer system – both software and hardware – that hosts the transaction programs. The transactional processing system is structured in a special way. This has some several components like end-user, Front-end program, Request Controller, Transaction server, Data base system (Bernstein & Newcomer, 2009)

The OLTP systems (Online Transactional Processing systems) are useful for addressing the operational data needs of the organization. They cannot support the business manager’s queries for decision support. The Data warehouse queries involve analytics involving aggregation, drill-down and slicing and dicing of the data that are supported by the OLAP systems (Sen & Sinha, 2005)

The transactional processing systems processes the data in their raw states as it arrives. Data warehouse systems integrate data from multiple source systems into a database suitable for querying. Data warehouse systems execute two types of workloads. They are a batch work load to extract data from sources, cleaning the data to reconcile discrepancies among them, transform them into a common shape so that it can be loaded into data warehouse; and queries against the data warehouse. They can range from small requests to very complex queries that generate complex reports. The transactional processing systems on the other hand involve some short updates and queries (Bernstein & Newcomer, 2009)
<table>
<thead>
<tr>
<th></th>
<th><strong>Transaction Processing</strong></th>
<th><strong>Data warehouse</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Isolation</td>
<td>Serializable, Multi Programmed Execution</td>
<td>No Transaction Concepts</td>
</tr>
<tr>
<td>Work Load</td>
<td>High Variance</td>
<td>Predictable Loading and High Variance queries</td>
</tr>
<tr>
<td>Performance Metric</td>
<td>Response time and throughput</td>
<td>Throughput for loading and response time for queries</td>
</tr>
<tr>
<td>Input</td>
<td>Network of display devices submitting requests</td>
<td>Network of display devices submitting queries</td>
</tr>
<tr>
<td>Data Access</td>
<td>Random Access</td>
<td>Possibly sorted for loading, Unconstrained for queries</td>
</tr>
<tr>
<td>Recovery</td>
<td>After failure, Ensure database has committed updates and no other</td>
<td>Applications responsibility</td>
</tr>
</tbody>
</table>

Table 3 Transactional systems vs. Data warehouse systems Characteristics.

ERP systems are not Decision support systems but they will greatly leverage the capability of decision making in the organizations using the data in them. ERP systems are the typical transactional processing systems that will enable the flow of the information from all the functional units of the business systems (Power & Sharda, 2009). They are mainly focused in the transactional processing of the data and weak on analytics. The BI systems are mainly focused on the analytics part of the business.

The development of BI projects is fundamentally different from transactional processing systems development. These projects are data-driven Business Integration Projects. Here the focus is data (Jim, Larissa, Chris, & Wyatt, 2009).

Within the last few decades there has been an extensive discussion over the systems that can support the decision making. Those systems are called as decision support systems in Information system literature. These broad classes of DSS (decision support systems) are labeled as MSS (management support systems). Now there is a common agreement on the definition of DSS and what it constitutes (Ariav & Ginzberg, 1985; Clark, Jones, & Armstrong, 2007; Maguire, 1978).

The decision making support systems that have development, design and research premises that are built around the problem situation are known to be KMS (Knowledge management systems) and BI (BI) systems. These two technologies have been central in improving the qualitative and quantitative knowledge available to the decision makers (Clark, et al., 2007; Cody, Kreulen, Krishna, & Spangler, 2002).

BI is a broad category of application programs and technologies for gathering, storing, analyzing, and providing access to data to help enterprise users make better business decisions. BI applications include the activities of decision support, query and
reporting, online analytical processing (OLAP), statistical analysis, forecasting, and data mining (Rossetti November 2006).

BI simplifies information discovery and analysis, making it possible for decision makers at all levels of an organization to more easily access, understand, analyze, collaborate, and act on information, anytime and anywhere (Hanumat, Venkatadri, & Manjunath, 2010; Microsoft, 2008)

BI can also be termed as competitive intelligence which is both a process and a product. As a process it is a set of methods for achieving the success in the global environment and as a product it gives information about the competitors activities from private and public sources (Vedder, Vanecek, Guynes, & Cappel, 1999)

The term BI was first coined by Howard Dressner an analyst from Gartner group in the year 1990. BI is especially used in the practice, in the world of analytics. According to the Gartner report in a survey of 1400 CIO’s (Chief Information Officers), majority opined that BI is the chief technology priority of the businesses for 2007 (Forsling, 2007)

The significance of BI is vibrant in both research and Industry. Data warehouses of sizes 10s to 100s of terabytes have become so common. With the proliferation of hardware and software technologies their capabilities have extended to a great deal (Surajit Chaudhuri, 2011)

BI is primarily involved in two main activities they are getting data in and getting data out. Getting data is traditionally referred to as Data Warehousing that is gathering data from a set of source systems to an integrated data warehouse. Data warehousing team extracts data from various sources of data and transforms into meaningful data for decision support. The getting of data is the second primary activity. The second activity is commonly called as BI by which the organization can fully realize the usage of the data warehouse developed (Watson & Wixom, 2007)

The typical BI project is guided by the following goals.

- Orientation towards business opportunities rather than transactional needs
- Implementation of strategical decisions, not only departmental or operational decisions
- Analysis based on business needs, which is the most important of the process
- Cyclical development process, focused on evaluation and improvement of success (Ion C. Lungu 2005)

According to “Olszak & Ziemba” building and Implementing BI systems involves two main stages. They are

1. Creation of BI
2. Use (Consumption) of BI
The creation of BI consists of the following activities.
   a. definition of the BI undertaking, i.e. determination of the BI system development strategies
   b. Identification and preparation of source data
   c. Selection of BI tools
   d. Designing and implementing of BI
   e. Discovering and exploring new informational needs and other business applications and practices.

The usage of BI consists of the following activities
   a. logistic analyses that enable to identify partners of supply chain quickly
   b. Access, monitoring and analyses of facts
   c. Development of alternative decisions
   d. Division and co-operation
   e. Change in the effect of company performance (Olszak & Ziemba, 2007)

![Figure 4 BI Framework "Getting data in & getting data out"(Watson & Wixom, 2007)](image)

The data warehouses typically are developed by one of the two methods proposed by either Ralph Kimball or Bill Inmon without referring to any of the development methods that are used for traditional system developments (Goede & Huisman, 2010)

Data warehousing methodologies share a common set of tasks they are Business Requirements Analysis, Data design, Architecture design, Implementation and deployment. Initially the project starts with the gathering of the requirements using either interviews or brainstorming or requirements elicitation techniques. Then after this a very high-level conceptual data model design is done. This is mainly done by two data modeling techniques. They are Entity-Relationship model or Dimensional Modeling technique that has Fact tables and Dimensions. The data warehouse architecture can be broadly classified into two types. They are Enterprise Data warehouse design and Data mart design (Sen & Sinha, 2005)
A typical Data warehouse project has the following critical steps or processes. They are Business Requirements definition, Data Acquisition, Architecture, Data Quality, Warehouse Administration, Metadata Management, and Data Access, Database design and Build, Documentation, Testing, Training, Transition, Post-Implementation Support (Collier, 2011)

The Data warehouse architecture requires the following set of discrete technical skills. They are (Collier, 2011)

1. Data Modeling
2. ETL development
3. Data Cleansing
4. OLAP design
5. Application development
6. Production Automation
7. General systems and Data base administration

The Business intelligence systems architecture is mainly based on the three levels. They are

First Level: Data Management

At this level the Data warehousing resides and the various data sources are gathered and connected together. The data warehouses can be used for analyzing or either creating reports. It can also be implemented without a data warehouse that is the data can be directly analyzed by getting data from data sources but that is a tedious process.

Second Level: Model Management

This is about building logical and physical models that can be used for statistical interpretation and analysis and forecasting.

Third Level: Data Visualization tools

This provides a visual drill down capacity to visualize data graphically and to analyze the complex relationships by the examiners.

The BI systems development lifecycle consists of the same steps and phases as that of the traditional transactional systems: Pre-study Phase, Project Planning, Analysis, Design, Construction and Implementation.

These are the major differences between the BI applications and Stand-alone Applications (Ion C. Lungu 2005) They are
BI applications vs. Standalone applications

<table>
<thead>
<tr>
<th>BI applications vs. Standalone applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Business opportunity</td>
</tr>
<tr>
<td>2. Implement across the organization</td>
</tr>
<tr>
<td>3. Strategic Information Requirements</td>
</tr>
<tr>
<td>4. Best deployed as Release/evaluate environment</td>
</tr>
<tr>
<td>1. Business Needs</td>
</tr>
<tr>
<td>2. Department decision support</td>
</tr>
<tr>
<td>3. Operational Functional Requirements</td>
</tr>
<tr>
<td>4. Best released at the same time with all functional capabilities</td>
</tr>
</tbody>
</table>

Developing a DW/BI system is fundamentally different from developing application software. The data warehouse/BI projects involves data integration efforts like data standardization, enterprise data modeling, business rules ratification by major business stakeholders, coordinated ETL data staging, common meta data, collectively architected (designed) databases, and so on. These are not specific to standalone transactional systems (Moss, 2009).

But for BIS development lifecycle these phases have several steps in them (Liang and Miranda 2001; Power and Sharda 2009).

<table>
<thead>
<tr>
<th>BI project Life cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Justification</td>
</tr>
<tr>
<td>Step 1: Business case Assessment</td>
</tr>
<tr>
<td>2. Planning</td>
</tr>
<tr>
<td>Step 2: Enterprise Infrastructure Planning</td>
</tr>
<tr>
<td>Step 3: Project Planning</td>
</tr>
<tr>
<td>3. Business Analysis</td>
</tr>
<tr>
<td>Step 4: Defining Business needs and project requirements</td>
</tr>
<tr>
<td>Step 5: Data Analysis</td>
</tr>
<tr>
<td>Step 6: Application Prototyping</td>
</tr>
<tr>
<td>Step 7: Meta Data Analysis</td>
</tr>
<tr>
<td>4. System Design</td>
</tr>
<tr>
<td>Step 8: Data Design</td>
</tr>
<tr>
<td>Step 9: Designing ETL process (Extract/Transform/Load).</td>
</tr>
<tr>
<td>Step 10: Meta Data Repository Design</td>
</tr>
<tr>
<td>5. Construction</td>
</tr>
<tr>
<td>Step 11: ETL development</td>
</tr>
<tr>
<td>Step 12: Application Development</td>
</tr>
<tr>
<td>Step 13: Data Mining</td>
</tr>
</tbody>
</table>
Step 14: Developing Meta Data Repository

6. System Development
Step 15: Implementation
Step 16: Release Evaluation

Table 5 BI development stages (Ion C. Lungu, 2005)

A typical BI architecture looks like above. It includes extraction, transformation, loading of data from various source systems into a centralized data warehouse from which cubes can be developed that are normalized for advanced analytics through multidimensional view. The results are often presented in reports with charts, graphs, diagrams to the users (R. Kimball, 2008; Olszak & Ziemba, 2007).

Generally while deciding on the style of the architecture for the data warehouse there are mainly two different styles one is “Bill Inmon style” and the other is “Ralph Kimball Style”. Bill Inmon is considered to be the father of data warehousing which follows 3rd normal form format for extraction and transformation of data. Ralph Kimball style follows Dimension & fact arrangement for data (R. Kimball, 2008; Kirkwood, 1998).
The first part of the BI revolves around the design and implementation of the data warehouse. The second part of the data warehouse is mostly concentrated on writing code to extract the data from the data warehouse to analyze the data (Jim, et al., 2009).

### 3.5.2 Case for Agile BI development

Forrester defines agile BI as an approach that combines processes, methodologies, tools and technologies, while incorporating organizational structure, in order to help strategic, tactical and operational decision-makers to be more flexible and more responsive to ever-changing business and regulatory requirements (Violino, 2011).

Most of the data warehousing initiatives need continuous incremental additions to their data warehouses. Data warehouses are perceived to be on-going systems rather one time done systems (Anahory & Murray, 1997).

The development of BI projects should be planned in a way that it produces some functionality or delivers a capability initially so that it sets up a platform for the future enhancements. So the idea of solving the problems all in one big shot is not possible. The development of BI project should be developed in an Iterative approach that delivers value to the organization in Increments (Mike, 2008).

BI development is often incremental in nature. Unlike transactional processing system like OLTP there is always a need for the BI systems to change for improvements. As the BI systems give strategic support for taking management decisions they should continuously change in accordance with the organizations internal requirements. For this BI development there need to be very good plan before starting the BI project (Ko & Abdullaev, 2007).

There is very high need for the standard that promotes the coherent team work. BI projects are often organization centric or specific. So a BI system is developed for the whole organization. The BI system should be developed with the emphasis on business perspective rather than the technological perspective. A BI system typically should consist of three applications they are, ETL services, Data warehouse & Integrated OLAP and Front End Tools (Ko & Abdullaev, 2007).

There is no reason why we can’t deliver value to our end users in weeks, days or hours rather than months. The agile methodology for BI can impact greatly on how users perceive the outcome from BI in a positive way (Bruni, 2011b).

There is a greater focus shift towards agile from traditional waterfall methodologies. The popular agile approach to the software development has caught the attention of the data warehouse community. (Dan, 2010) The application of agile principles to BI is natural as the developmental nature of BI systems is iterative and incremental (Larson, 2009).

Agile ideals fit into the BI world very well. But the successful application of the agile principles depends on understanding how they can be applied and also keeping focus on the information results rather than the software itself (Larson, 2009).
3.6 Agile system development

3.6.1 Need for Agile Methods

Software development methodology is a framework that is used to structure, plan, and control the process of developing an information system - this includes the pre-definition of specific deliverables and artifacts that are created and completed by a project team to develop or maintain an application (Elliott, 2004).

Software development has become always a complicated task. With the advent of the new technologies, tools and techniques there is always a demand for the software that exceeds the skill of the established software developers. As there is always a higher imagination or expectation from the customer side than what the current resources and technologies we have. We constantly need new ways for achieving them. We need to incorporate change as an attribute to the system that is being developed (Barry & Lang, 2003).

According to (Nandhakumar J, 1999) there are several software development approaches that are being proposed but only few are surviving. Information system development often brings in several vendor-specific techniques, tools and methodologies. Each is specialized for certain processes of IS development. The traditional development methodologies are having less control over the development of information system and are also they appear to be too mechanistic in-order to follow in detail. The agile methods evolved from the inability of the plan driven methodologies to handle the rapidly changing environments (Jim Highsmith, 2002).

According to QSM associates Michael Mah more than 200 software development projects couldn’t find their initial plan to measure with. This indicates that the primary goal of the software development project is not to stick to the plan any more. Today’s software development main goal is to give customer satisfaction. Satisfying the customer at the software delivery is more important than satisfying them at the beginning. So earlier it is believed to be with careful documentation of the requirements at the beginning can deliver the software in time by keeping away the potential changes that come in future. But if the developer is not responding to the customer then the provider is not providing necessary features to the software (J. Highsmith & Cockburn, 2001). This condition of software development methodologies which are reluctant to change during the course development led to the evolution of the agile methodologies. With this as the background we move towards the agile methodologies.

The term agile is first coined at the first agile software development alliance in 2001. The word agility means to move quickly and easily according to Oxford dictionary. Agile methods help the organization or a team to adopt the subset of principles and practices based on the Culture, Values and type of systems that they developed. Every organization has its own tailored made agile methods based on their needs. Agile methods are helpful for customization. Organizations adopt agile methods based on the culture, values and the systems they produce. Every agile method has an objective that they intend to achieve using them and the principles that
assist them to reach those objectives and the practices that are reflective on the principles it is based on (Soundararajan, 2011)

The introduction of the agile manifesto in just less than a decade ago led the software industry to undergo some tremendous changes. Since the development of Information systems is commenced there are a number of tools, techniques and methods have been proposed.

With the agile principles being enunciated in the agile manifesto there are a variety of agile methods that are proposed in relation to the degree of adherence to the agile principles. They include Extreme programming, Scrum, Lean software development, feature driven development and crystal methodologies (Dingsøyr, Nerur, Balijepally, & Moe, 2012)

Agile methods have become the nerve of the software development industry. With the growing impact there are many to support and embrace agile methods and also many who criticize them. There is an effort from some group to mix agile and plan driven approaches. With growing significance of the agile methodologies the researchers have found that even the plan-driven methodologies have their own home grounds at which they can perform with better efficiency. Hence many researchers are devising plans to mix or bridge between the agile and plan driven methods for maximizing the productivity of software development (Boehm & Turner, 2003)

Many development methods that are defined at that time started with a complete documentation of the requirements at the initial system development. But business and technical requirements often change with the ongoing time. So it has become obvious that projects got delayed due to the embracing of the change. So by then practitioners started to develop new methods to embrace and accommodate change rather than rejecting them. These methodologies are developed in 3 continents simultaneously. They are In Europe they developed DSDM (Dynamic system development methodology), Feature driven development in Australia, and Extreme programming, Crystal, Adaptive software development and Scrum in USA. They rely basically on the same agile principles but each practitioner authored them individually (Williams & Cockburn, 2003)

The Manufacturing industry underwent some radical changes with the introduction of the lean production system by Toyota. The traditional production systems based on the Ford and Taylor are challenged by the Toyota new system (i.e. Lean Production) which is extremely successful. The main success of the system is attributed to their ability to respond to change i.e. flexibility along with reducing waste (Schmidt & Lyle, 2010)

Core principles and Values of Agile

In February 2001, at the Lodge of resort and ski in the mountains of Utah 17 researchers met to talk, relax and find the common ground. This resulted in the agile software development alliance. They categorized each of their methodologies as light weight methodologies. Their common goal is to provide customer satisfaction and a product of high quality. They coined the term agile. They have these four values as the underlying values for the agile position(Jim Highsmith, 2002).
They are

1. Individuals and Interactions over Processes & tools.
2. Working software over comprehensive documentation.
3. Customer collaboration over contract negotiation.
4. Responding to change over following a plan.

These are the four comparative values which recognize the importance of the terms that are in the right but assert the dominance of the terms in the Left (Cunningham, 2001).

The above principles enable to deal with the change in the requirements at any point of system development process.

Any agile method is governed by the above four principles to make the software development process more flexible and thereby making it more successful.

### 3.6.2 Agile Methodologies

Agile methods main focus is on early releases of the software using the collaborative methods like pair programming and refactoring etc. with the customers or clients on the development site. The development team releases working software rather than the prototypes to the customers in every release (Reifer, 2002)

On a survey performed by Reifer on firms using agile methods this is the list of practices that they considered agile practices: collective ownership, concurrent development, continuous integration, customer collaboration, daily standup meetings, product demos instead of documents, Extreme Programming (XP), frequent product releases, full stakeholder participation, individuals and interactions, just-in-time requirements, metaphors instead of architectures, nightly product builds, pair programming, rapid application development, refactoring, retrospectives, stories for requirements, team programming, and test-driven development (Reifer, 2002)

There is no common agreed definition of what agility means. According to Qumer & Henderson “Agility is a persistent behavior or ability of a sensitive entity that exhibits flexibility to accommodate expected or unexpected changes rapidly, follows the shortest time span, uses economical, simple and quality instruments in a dynamic environment and applies updated prior knowledge and experience to learn from the internal and external environment.” (Qumer & Henderson-Sellers, 2008)

When we apply the above definition to the software development methodology the following definition is arrived for the agile methods. It is as follows “A software development method is said to be an agile software development method when a method is people focused, communications-oriented, flexible (ready to adapt to expected or unexpected change at any time), speedy (encourages rapid and iterative development of the product in small releases), lean (focuses on shortening timeframe and cost and on improved quality), responsive (reacts appropriately to expected and
unexpected changes), and learning (focuses on improvement during and after product development)’’ (Qumer & Henderson-Sellers, 2008)

Most of the methods that are considered to be agile have the following topics as common. They are Change, Planning, Communication & learning (Koch, 2004)

There are at least 13 agile methods that are under practice. The most used being the Extreme programming and Scrum the following are also widely used like Feature driven development, Crystal methods, Dynamic systems development methods, Lean method, adaptive software development (Dybå & Dingsøyr, 2008; Strode, Huff, Hope, & Link, 2012)

These are the major agile methods that are being employed now. There has been considerable amount of literature reviewing the existing agile methods and contrasting them with each other (Pekka Abrahamsson, Oza, & Siponen, 2010; P. Abrahamsson, Salo, Ronkainen, & Warsta, 2002; Jim Highsmith, 2002; Koch, 2004).

They are

1. Scrum
2. Dynamic System Development Method (DSDM)
3. Crystal Methods
4. Feature-Driven Development (FDD)
5. Lean Development (LD)
6. Extreme Programming (XP)
7. Adaptive Software Development (ASD)
8. Agile Modeling

**Scrum**

Scrum is a project management framework which focusses on software development in 30-day sprint cycles in which a specified set of Backlog features are delivered. Scrum is named after scrum in Rugby. It was initially developed by Ken Schwaber and Jeff Sutherland in 1990’s and with latter collaborations Mike Beedle joined. The core practice in scrum is the 15-minute meeting daily for co-ordination and integration. Scrum is being used to deliver several projects successfully (Schwaber & Beedle, 2002).

Scrum is more on simplification of the project management. It is rough simplification of the process based on iterative approach. It mainly consists of the three roles, three documents and three meetings. They are Roles: product owner, team and scrum master. Three documents: Product backlog, sprint backlog & sprint results. Three meetings: sprint planning meeting, daily scrum meetings, sprint review.

Test driven development: Automation of test cases is crucial for successful development of the software. Writing manual test cases for each and every iteration and sprint is laborious task. So developing test cases immediately after design of the iteration is done and testing it before implementing it is very helpful (Hansmann, 2009)
Dynamic Systems Development Method (DSDM)

The Dynamic Systems Development Method was developed in the U.K. in mid 1990s. It is an outgrowth and extension of Rapid Application Development practices. DSDM has the best supported training and documentation than any other agile software development methods at least in Europe. There are nine principles that include Active User involvement, frequent delivery, team decision making, integrated testing throughout the project lifecycle, and reversible changes in development (Stapleton, 1997)

Crystal Methods

Alistair Cockburn is the author of the “Crystal” family of people-centered methods. Alistair is considered as a “methodology archeologist” who has interviewed dozens of project teams worldwide trying to separate what actually works from what people say should work. Alistair and Crystal, focuses on the people aspects of the development-collaboration, good citizenship and cooperation. Alistair uses project size, critically and objectives to craft appropriately configured practices for each member of the crystal family of methodologies (Cockburn, 2002)

Feature-Driven Development

Jeff De Luca and Peter Coad collaborated on Feature-Driven Development. FDD consists of a minimalist, five-step process that focuses on developing an overall “Shape” object model, building a feature list, and then planning-by-feature followed by iterative design-by-feature and build-by-feature steps. FDD’s processes are brief (each is described on a single page), and two key roles in FDD are chief architect and chief programmer. FDD differs from XP in its “light” up-front architectural modeling (Palmer & Felsing, 2001)

Lean Development

The most strategic oriented is also the least known: Bob Charette’s Lean Development, which is derived from the principles of Lean production, the restructuring of the Japanese automobile manufacturing industry that occurred in the 1980’s. In LD, Bob extends traditional methodology view that the change as a risk of loss to be controlled with restrictive management practices to a view of change is as producing “opportunities” to be pursued using “risk entrepreneurship”. LD has been used by a number of large telecommunications projects in Europe (Schmidt & Lyle, 2010)

Extreme Programming

XP or Extreme programming was developed by Kent Beck, Ward Cunningham, and Ron Jeffries. XP core values are Community, Simplicity, Feedback, and Courage. Important aspects of XP are in contributing to the altering the view of the cost of change and its emphasis on technical excellence through refactoring and test-fire development. XP provides a system of dynamic practices, whose integrity as a holistic
unit has been proven. XP has clearly garnered the most interest of any of the Agile approaches (Beck, 1999; Jim Highsmith, 2002)

**Adaptive software development**

Adaptive software development has the philosophical background for the agile methods. It shows how organizations can respond to the turbulence of the current business climate by harnessing rather than avoiding change. This consists of iterative development, feature-based planning, customer focus group reviews and an agile management philosophy called Leadership-Collaboration management (Jim Highsmith, 2002)

**Agile Modeling**

In contemporary software development, the modern development processes do not provide the detailed requirements up in front nor do they focus on detailed models. They evolve their models on time with the changing requirements of their stakeholders. Evolutionary (incremental and iterative) development is a norm for the software development in the current software industry. Agile Data method defines a collection of philosophies and techniques for development of databases using the evolutionary approach.

Most software development processes like extreme programming (XP), RUP (rational unified processes) and feature driven development work on evolutionary approach (i.e. incremental and iterative approach). They are adopting some evolutionary techniques like

1. Agile Model Driven Development
2. Database Refactoring
3. Evolutionary performance tuning
4. Test-driven design (TDD)
5. Configuration management of database artifacts
6. Developer sandboxes
<table>
<thead>
<tr>
<th>ASDM</th>
<th>Origin</th>
<th>Principles</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSDM</td>
<td>DSDM Consortium in 1994</td>
<td>Keep time and resources fixed while adjusting functionality according to requirements.</td>
</tr>
<tr>
<td>XP</td>
<td>Kent Beck in 1996 (Hislop et al. 2002)</td>
<td>4 Values: Communication, simplicity, feedback, courage. 4 Basic activities: Coding, testing, listing, designing.</td>
</tr>
<tr>
<td>FDD</td>
<td>De Luca and Coad in 1997 (De Luca 2005)</td>
<td>Built around a core set of “best practices”. Business practices decomposed into features. Features are small items useful for users.</td>
</tr>
<tr>
<td>Crystal</td>
<td>Cockburn in 1999 (Cockburn 2001)</td>
<td>Address different kinds of project requirements with different kinds of Crystal ASDMs. Family of Crystal ASDM organized in terms of system criticality and number of people involved. Focus primarily on communication</td>
</tr>
<tr>
<td>LD</td>
<td>Carette in 1980 but popularized in 2003 by Poppendieck (Poppendieck 2003)</td>
<td>Complete the project in one-third of the time, one-third of the budget, one-third of the defect rate. Management philosophy rather than a development process.</td>
</tr>
<tr>
<td>AM</td>
<td>Agile modeling by Scott Ambler in 2002 (Scott Ambler, 2002)</td>
<td>Light weight modeling and documenting of software systems.</td>
</tr>
</tbody>
</table>

Table 6 The main characteristics of each one of these eight ASDMs (Goede & Huisman, 2010)

3.7 Agile Strategies

For system development methodologies are largely influenced by the organizational culture. The organizational culture diverge over four dimensions they are management, technology, people & process. They are guided by the visions and values of the organizations (Vinekar, Slinkman, & Nerur, 2006)

System development methodologies are intended to be useful to the developers. The interesting thing observed is that the organizations are either using a methodology with some changes that are in-line to the organization context or not at all using. This is due to the reality that the methodologies are not being developed with the organizations as context (Fitzgerald, 1997)
At this point with a lot of agile methodologies and practices evolving it very essential to have a look on what is agility and what does it constitutes. The current knowledge of the body of agile lacks the clarity of what is it that should be considered as agility. Agility in system development methodologies at organizations should be multifaceted and contextual in nature. The nature of agility is subjective in nature that is it depends on the environment of the project. So every organization should define its own version of what agility is based upon its project and organizational environment (Lyytinen & Rose, 2006)

So with this in context let’s move forward and see what the IT Company considers agility and what practices or methods they adopted based on its projects and organizational culture.

**Agile strategies used for Application development and maintenance at an IT company which is considered for this study:**

The IT Company is moving forward in focusing more on Agile and Lean for Application development and maintenance on Value creation and Value Delivery.

**Value Creation:** Finding out the need of the customer that is what they are really in need of and expecting from us and then creating it.

**Value Delivery:** Giving our customers what we have promised and ensuring that they are satisfied.

**Agile** software development is a group of software development methodologies based on iterative and incremental development, where requirements and solutions evolve through collaborative between self-organizing, cross-functional teams. It promotes adaptive planning, evolutionary development and delivery, a time-boxed approach, and encourages rapid and flexible response to change. The key element is to deliver prioritized value to the business early, and often, in the process.

**Lean** development can be summarized by seven principles, very close in concept to the lean manufacturing principles adapted from the Toyota Production Systems:

- Eliminate waste
- Amplify learning
- Decide as late as possible
- Deliver as fast as possible
- Empower the team
- Build integrity in
- See the whole

**Scrum** is the promoted way of working in the cases of Projects where there is a more stable / "boxed" development environment which is different to Maintenance where there are bug fixes and change requests to throw into the equation.

**Kanban** will be the default way of working in the case of Maintenance assignments to give us greater degree of flexibility to respond to changing priorities where there
are bug fixes and change requests to deal with.

Both Scrum and Kanban adopt Agile and Lean principles at their core and create value for Customers (Wilson, 2012)

These are the agile strategies that are used by the IT Company.

1. Lean methodology/ Kanban
2. Agile/Scrum/XP practices

**Agile Practices:**

In addition to implementing Scrum and Kanban we will have 13 supporting Agile "practices". Some of these will be seen as core practices and to be used in all cases and others more as tools to use depending on the circumstance.

The Agile “practices” that we identified apart from Scrum and kanban are as follows:

**Securing the solution**

1. Requirements Breakdown
   Ensures the architecture fulfills the Requirements
2. Domain Modeling
   Ensures Business & IT uses the same language
3. Prototyping
   Visualizes the requirements & workflows

**Final Scale feedback**

1. Pair Programming
2. Planning Game
3. Test-driven development
4. Whole Team

**Continuous Process**

1. Continuous Integration
2. Refactoring
3. Small Release

**Shared Understanding**

1. Coding standard
2. Simple design
3. System metaphor
4. Collective code ownership

**Programmer Welfare**

1. Sustainable Pace
These are some of the Agile Principles behind the agile manifesto.

1. Welcoming changing requirements, even late in the development
2. Delivering working software frequently
3. Business people and developers work together daily throughout the project
4. Build projects around motivated individuals
5. Continuous attention towards technical excellence and good design
6. Meet the customer needs through early and continuous delivery of valuable software

All most all of the agile methodologies primary goal is to write and deliver the software code as early as possible.

**Agile/Lean way of working at IT Company:**

![Extreme Programming (XP)](image)

Figure 6 Agile/Lean way of working Agile DW/BI

There are mainly two different methods for the development of the data warehouses. They are Bill Inmon style and Ralph Kimball Style. The Bill Inmon methodology is primarily driven by the data, making it data driven or data centric. The ralph Kimball style is requirements driven methodology (Goede & Huisman, 2010)

The Ralph Kimball life cycle methodology encourages iterative development, dimensional modeling with maximum involvement of the user. These ideas by Kimball easily match with the agile system development methodologies core principles (Ralph Kimball, Ross, Thornthwaite, Mundy, & Becker, 2011)

Hence the Ralph Kimball style has much to adopt from the agile system development methodologies to bring in agility for the data warehouse development. Hence all these practices are discussed with Ralph Kimball development methodology in perspective.
There are some best practices suggested by several authors for bringing agility into DW/BI projects. They include technical, project management, architectural and user collaboration practices.

The primary objective to introduce the agile to DW/BI systems is to deliver frequently a high-quality, high value, working DW/BI system responding to change (Dan, 2010)

Requirements gathering phase is of utmost importance in a DW/BI project as capturing the accurate business requirements and presenting it to the IT team is crucial. This paves the way to have a look at the availability of the data sources and quality of the data. From this depends of what can and cannot be delivered to the business team. Based on the outcome of the Requirements gathering phase the future increments are planned (Larson, 2009)

The below are the various practices recommended by various authors for Agile BI/DW development (Collier, 2011; Dan, 2010; Larson, 2009; W.Ampler, 2011)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Agile Project Management</td>
<td>1. Evolving Excellent Design</td>
<td>1. Do some initial architecture envisioning i.e. scalable architecture</td>
</tr>
<tr>
<td>2. User stories for BI systems (Requirements gathering)</td>
<td>2. Test-Driven Data warehouse development</td>
<td>2. Model the details just in time (JIT).</td>
</tr>
<tr>
<td>4. Value driven development or delivering working software regularly</td>
<td>4. Project Automation</td>
<td>4. Focus on usage</td>
</tr>
<tr>
<td>5. Production quality</td>
<td></td>
<td>5. Don't get hung up on &quot;the one truth&quot;</td>
</tr>
<tr>
<td>6. Barely sufficient processes</td>
<td></td>
<td>6. Organize your work by requirements</td>
</tr>
<tr>
<td>7. Testing throughout the lifecycle and Automate as much as possible. Note: Regression testing should be completed for each change done to the system.</td>
<td></td>
<td>7. Active stakeholder participation</td>
</tr>
<tr>
<td>8. Collaboration most preferably face to face communication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Self-organizing and self-managing team will help in emerging best requirements, architecture &amp; design.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Strive for iterations of 1 to 2 weeks</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7 Various practices for BI/DW development

More practices:

1. Use good tools to speed up development in design and automate as much as you can.
2. Adopt common development standards.
3. Don't underestimate legacy data challenges.
4. Travel light.
5. Adopt a lean approach to data governance
6. We should have a formal process for deployment and maintenance of the application.
7. Data validation and verification for iteration should be done.
8. Iterations should be time boxed.

Graziano formulated the following 12 principles of agile methodologies for applying to the Data warehouses. They are the following (Goede & Huisman, 2010)

<table>
<thead>
<tr>
<th>Principle</th>
<th>Application in DW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Satisfy the customer through early and continuous delivery of valuable software</td>
<td>Valuable software includes BI reports and Dashboard interfaces</td>
</tr>
<tr>
<td>2. Welcome changing requirements even late in the development</td>
<td>Use code generators; start with normalized tables.</td>
</tr>
<tr>
<td>3. Deliver working software frequently</td>
<td>Deliver one subject area at a time.</td>
</tr>
<tr>
<td>4. Business people and developers must work together daily throughout the project.</td>
<td>Good idea, but politics and priorities may interfere.</td>
</tr>
<tr>
<td>5. Build projects around motivated individuals.</td>
<td>Keep units of small work; Don’t deliver a single Enterprise DW</td>
</tr>
<tr>
<td>6. Use face-to-face communication inside development</td>
<td>Use daily meetings and keep documentation efficient.</td>
</tr>
<tr>
<td>7. Working software is the primary measure of progress.</td>
<td>Working BI reports and ETL software is important.</td>
</tr>
<tr>
<td>8. Sustainable development at constant pace.</td>
<td>Keep units of work small, Use XP and Pair Programming</td>
</tr>
<tr>
<td>9. Continuous attention towards Excellence and good design enhances agility.</td>
<td>Frequent design reviews leads to faster delivery.</td>
</tr>
<tr>
<td>10. Simplicity- the art of maximizing the amount of work not done-is essential.</td>
<td>Use metadata effectively; use code generators; online data dictionary.</td>
</tr>
<tr>
<td>11. The best architectures, requirements, and designs emerge from self-organizing teams.</td>
<td>Don’t micros manage staff, Have shared responsibilities.</td>
</tr>
<tr>
<td>12. Team should reflect at regular intervals how to become more effective, and should adjust accordingly.</td>
<td>Use self-organizing teams with maximum buy-in to the solution.</td>
</tr>
</tbody>
</table>

Table 8 Application of ASDM principles in DW adapted (Goede & Huisman, 2010)

These are the six architectural principles that are proposed by Lee Arnett.

1. Assemble a staff conditioned to change
2. Use metadata-driven development tools
3. Isolate subjects and data tables
4. Bit-size analysis through denormalization
5. Use surrogate keys
6. Frequent source system feeds (Arnett, 2002)

Agile BI development is about adopting the best practices of agile software development and also with adapting the DW/BI’s own best practices for agility. Agile software development main goal is to deliver the software early, frequent and sustainable delivery of working software that gives business value. In Agile
development the focus is on iterative and incremental development which suits the nature of BI requirements as they change frequently. These are agile practices that are being considered for BI now (James E. Powell, 2012)

1. Adopt agile project management.
2. Development practices such as scrum and XP.
3. The use of Test Automation.
4. Reusable design patterns.

Agile modeling helps a lot as it allows the collaboration with the BI stakeholders in designing the database there by gather their feedback early rather than waiting for the less direct requirements gathering techniques like Decode data requirements from interview notes, requirement documents or user stories etc. (James E. Powell, 2012)

3.8 Summary of theoretical findings

The summary of the theoretical findings will help us to answer the main Research question.

Q. How to adopt agile way of working for BI development?

This Main research question is answered by answering the two sub-research questions under the main research question. They are answered in two subsections.

3.8.1 Typical transactional processing systems development vs. the BI assignment development

A Typical BI project aim and focus differs from that of the Transactions systems as follows.

- Orientation towards business opportunities rather than transactional needs
- Implementation of strategical decisions, not only departmental or operational decisions.
- Analysis based on business needs, which is the most important of the process
- Cyclical development process, focused on evaluation and improvement of success

The strategies used for the transactional processing systems vs. the BI assignment development are quite similar. They follow the same development phases like Requirements Gathering, Design, Development, Implementation and Maintenance. The same set of strategies can be used for both the transactional systems development and the Business intelligence development projects.

The usage of the agile strategies can be much easier in transactional systems development as the requirement doesn’t change as frequent as the requirements in the BI projects. As the requirements or the intended functionality expected to fulfill the
business need in the transactional processing systems are explicitly mentioned at the initial stages of the project. This indicates that the requirements based on the deliverable functionality can be divided and scheduled into a product backlog.

Hence we can use the scrum/ XP methodologies for the development of the project.

There are a number of major differences between designing a warehouse database and an operational system database (OLTP). For the data warehouse, we need to reconsider the approach and thought processes:

- We must focus on designing for optimum query access. For the operational environment, we also optimize for data manipulation tasks.
- A data warehouse is developed increment by increment.
- The final result is a dynamic structure, rather than the static operational structure.
- The data warehouse stores historical data by time period.

The BI team faces a lot of challenges:

- It’s unlikely to know what queries are going to be made.
- Most queries must have high performance with quick response time.
- Managing the volume of data (incl. Backup and restore).
- Controlling and managing operating costs.

A BI project is not developing code for the projects. It is Business Data Integration Project. A BI project can be clearly divided into two parts

1) Data Management (Through Data Warehouse)
2) Data Delivery (Through BI application)

Data drives in the main DW/BI projects. The scope of the project should be prioritized based on the functional requirements of the BI project. The release planning should be done based on the data effort that is needed for the each functionality takes. Data Management is the key important job here that requires nearly 80% of the time and effort consumed by the BI project. Data effort involves the key functions such as data profiling, data modeling, data standardization, data integration, data dispute resolutions (involving users from different departments), meta data capture, meta data storage, data lineage, ETL architecture, and database architectures. All of these aspects must be considered while scoping for the BI project.

3.8.2 How do we implement agile strategies in BI development?

The reason for advocating agile methods for BI is to deliver working deliverable early and frequently. The system also should be able to respond to the change during the development as it is the core value in the agile manifesto.
The Data warehousing architecture is implemented using well defined iterative approach in the IT Company considered. The following is recommended in iteration approach.

The usage of agile strategies for Data warehousing involves with the result of highly iterative approach with collaboration between users, developers and stakeholders.

First of all for achieving agility there should be a change in mind set of all the team members. Customer community member’s main intention should be to take sufficient time to explore the newly completed features and give feedback about it. The management community should expect and embrace unexpected change and project risk and uncertainty. The technical team should learn to work in a way which includes lot of discipline and rigor. The project interface community should be ready for daily project involvement and shift in their roles for the overall success of the project (Highsmith 2002; Ambler 2003; Collier 2011).

Highsmith proposed an APM (Agile Project Management) framework which is based on Envision – Explore cycle. This encourages the frequent collaboration between users, stakeholders & developers throughout the project cycle.

Envision: This is about envisioning the output of the project and then planning how to implement this using highly iterative and incremental approach.

Explore: All the phases cannot be built by using traditional waterfall model steps like requirements, design, code, testing, and implementation. There need to be sometimes to explore certain things before going for phase development. They need to explore write some code derive one or two features test them and communicate with the user and find if we are in the right direction.

In this research document earlier we discussed about the various flavours of agile methods.

The following are some of the practices to be adopted for making the project more agile. They are

1. Just Enough Design
2. Synchronize daily
3. Time boxing everything
4. Collocating teams
5. Attention to technical debt
6. Plan to capacity and monitor velocity
7. Track daily progress
8. Monitor story completion not task time

One of the major reasons identified for the failure of the BI projects is the failure to collaborate. One failure is to collaborate with the customers and the other with the team members. Agile development calls for a regular and continuous interaction with these groups. The builders should be involved on daily basis and the users should be involved on weekly basis.
There are Doers, planners and consumers. Doers are the ones who actually develop the project. Planners are the ones who develop the plans and consumers are the ones who are going to use them. The Doers are directly involved in the project daily and the planners are involved in the project weekly and the consumers are involved in the project as many times in a month.

**Evolving Excellent Design:**

For a system to be able to quickly respond to change and frequently deliver new features requires excellent data models and system design. Evolutionary design is based on good conceptual model at the beginning, and it should be followed by continuous refactoring towards excellent design.

Evolutionary design involves the following key developer practices. They are

1. Database refactoring
2. Evolutionary data modelling
3. Database Regression testing
4. Configuration Management
5. Developer sandboxes

Evolutionary design involves in producing excellence by effective usage of the agile modelling, Database Refactoring and design patterns.

**Test-Driven data warehouse development:**

For delivering BI systems of high quality in every iteration or sprint requires testing the iteration before delivering. Integration of testing to the development process enables in the delivery of a BI solution of high quality. Since manual testing is not always easy to use when going for the highly iterative development *test automation* is done. Usage of integrated automation and their benefits are complemented by *Test-driven Development* (Ken, 2005)

For a system to be fully agile it is necessary that the routine manual processes to be automated. Hence DW/BI teams use automated continuous Integration and deployment and monitoring.

After automation and continuous integration is running at the end of the iteration we come to a point of releases. For increasing the agility to the releases the author is recommending the usage of the Small releases practice.

The typical data warehouse projects require too much consumption of resources and the time. Hence the organizations are looking towards the Agile BI development that can deliver the usable results in the initial iterations. The common system development methodologies like scrum and extreme programming are applied to the BI. Agile project management delivers a great deal to both the Business and IT. Optimization of the infrastructure for the agile development approach eases the development process for the Agile BI development. System being hosted in the cloud much more helps to cope up with the continuous Data sources changes. The Business
and the Executive sponsors should commit to the Agile development approach then only it can directly reflect on the Agile BI architecture (Bruni, 2011a)

The first step Business Case Assessment can be done using the help of the user stories. By using the user stories the Business Analysts get an idea of possible solutions and their cost and benefits. The next step is that of Planning regarding the organization infrastructure and their capabilities.

The success of the project also depends in the way it is controlled & managed. The key to a successful BI solution is the integration of the business, data and technical domains, i.e. to bridge the gap between the business view and the IT technical view.

There are several requirements gathering techniques that can be used along the agile methods for effective Requirements gathering. They generally concentrate on the communication with the customer for the evolution of the requirements, changing requirements, Prioritizing requirements and delivering the most important functionalities first. The exploration and capturing of the functional and non-functional requirements is well tackled by the Scrum by using the User-stories and index cards (Kavitha & Thomas) User stories provide a quick way for gathering and organizing requirements without the need for the detailed requirements analysis upfront. The user stories provide the essence of the functions that the users need in the BI system and keep the details apart. Stories are gathered collaboratively during Project chartering and then prioritized on a product backlog that is continuously groomed and maintained by the product owner (Collier, 2011)

The Pre-study phase after identifying a business need is followed by Iteration Planning that involves the actual development and deployment of a data warehouse using iteration approach.

Management Methods:

- Agile Project Management
- Community cutomers and collaboration
- User stories for BI systems
- Self-Organizing teams for boosting performance

Technical Methods:

- Evolving Excellent Design
- Test-Driven Data warehouse development
- Version control for data warehousing
- Project Automation(Collier, 2011)

These are some of the DW/BI Modeling Best practices for the actual development of the Data warehouse.

**DW/BI Modeling Best Practices**

1. Do some initial architecture envisioning
2. Model the details just in time (JIT).
3. Prove the architecture early
4. Focus on usage
5. Don't get hung up on "the one truth"
6. Organize your work by requirements.
7. Active stakeholder participation

**DW/BI Lifecycle Best Practices**

1. Take an evolutionary approach.
2. Embrace change
3. Deliver working software regularly
4. Strive for iterations of one to two weeks
5. Test throughout the lifecycle
6. Involve operations and support people early (S. Ambler, 2003)

These are the following agile strategies that are used by the IT Company.

**Agile/Lean way of working at an IT company:**

![Extreme Programming (XP)](image)

We are using Scrum (Agile Principles) for BI projects and Kanban (Lean Principles) for Maintenance projects. Apart from these two we use 15 powerful Extreme Programming Practices.

BI (BI) projects in an IT company selected as a case study follow a gated Methodology/Framework.

The BI project development phases follow a set of phases and steps in academia that is almost similar to the project models followed by the Industry. For adapting the agile strategies for the BI projects we are recommending the following.

1. The Customers should be involved in early stages of the development. The team should adopt an agile mindset for working in an agile way.
2. There should be a commitment from the Organization i.e. executive’s sponsors and Business sponsors for the agile development approach. That they all should have an agile mindset.

3. As recommended above the usage of user stories for capturing architectural and data requirements should be done along with business needs.

4. There are some efficient management methods which we can employ for using in BI projects to develop using agile development approach. As we know that the requirements in the BI projects are taken first and they evolve continuously with the time as the data systems become bigger in size. We need to have an efficient way for gathering requirements i.e. user stories.

   - Agile Project Management
   - Community cutomers and collaboration
   - User stories for BI systems
   - Self-Organizing teams for boosting performance

5. These are the technical recommendations for the developing using agile development approach.

   - Evolving Excellent Design
   - Test-Driven Data warehouse development
   - Version control for data warehousing
   - Project Automation

   And there some common ideas that we need to buy to be more agile like.

   - Adopt common development standards.
   - Use good tools.
   - Don't underestimate legacy data challenges.
   - Travel light.
   - Adopt a lean approach to data governance.

Usually the iterations are done in 90 – 120 days iteration. This is too long time to wait for some working results. Hence we are trying to see if we can fit the agile way of working (i.e.) for these iterations to enable frequent delivery of working software. We also are trying to view the XP practices if they can be used in these iterations.

3.9 Arguments for an Empirical Study

The research was started with the basic question about how to develop BI systems with agile strategies. For answering this question again the main question is divided into two sub-questions. By theoretical study the first two sub questions are answered and with the help of which the main research question can be answered. In Qualitative study the results of the theoretical study are again verified with the Empirical Study.
4 EMPIRICAL FINDINGS

4.1 Purpose
The main purpose of the Empirical study is to collect data which is useful in answering the Main research question on, “How to adopt agile way of working for BI development?” The empirical study helps to identify the knowledge that is necessary for developing the BI projects using the current agile strategies. By thus we can come up with suggestions to implement agile strategies in BI projects.

4.2 Sampling

Sampling is a process of selecting units (i.e. people, organizations) from a population of interest so that by studying the sample we may fairly generalize our results back to the population from which they were chosen (Trochim, 2006)

The company for the case study met most of the characteristics mentioned in the earlier chapter. The IT Company selected is a very big product based company that is inundated with lots of data. By using BI tools there is always a great necessity to go for analysis etc. to forecast future trends and serving reporting need etc. As the BI projects are often huge in big companies there is every need for frequent and early deliverables and the systems that could cope up with change. The IT Company is taken as case study that can unveil interesting results at the end that can be generalized.

4.2.1 The Respondents

The respondent’s that are chosen for the interviews is completely based on their knowledge and experience in the relevant field that is in interest for study. Each respondent have very good experience in the BI development across various organizations and now in very big IT Company which has nearly 5500 employees across the globe. They have worked in various development phases of the BI projects both small and big. The respondents are selected from each development phases of a BI project like Business analysts from Requirements phase, Architects and developers from Design and development, Test lead from testing phase and maintenance manager from maintenance phase and a project managers to give the complete picture. This helps me for the interview to get the picture of how practically can Agile BI be implemented.

4.3 The interviews

Interviewing is a popular way of gathering qualitative research data because it involves talking and talking is natural. Interviews do not presuppose any statistical knowledge, and persons to be interviewed are considered to be respondents who are normally at hand and willing (Griffee, 2005)There are few interviewing styles they are Structured Interview; Semi structured Interview & Unstructured Interview. The structured interview starts with a specific set of questions that the researcher is
interested in and it’s more like a survey. The semi-structured interview gives some flexibility regarding the questions, the interviewer is free to move between his interested research areas. The unstructured interview is more like a conversation flexible and unrestricted (Virginia, 2012)

The interviews that are used for this research are termed as semi-structured. These interviews are designed to have number of Interviewer questions in advance but such prepared questions are designed to be sufficiently open for the subsequent questions of the interviewer that cannot be planned in advance but must be improvised in a careful and theorized way (Wengraf, 2001)

The main reason for taking up this kind of interview is that this thesis is mainly based upon the qualitative research method, chapter 2.1. The main focus of the qualitative research method is to provide an in-depth understanding about a specific area of interest. A semi structured interview facilitates the interviewer to explore and expand his knowledge by communicating with the respondents and there creating a broad picture of the subject area that is under study.

The semi-structured interview provides the scope of exploration and also it enables the result of the interview to be structured and analyzed in a relatively easy cohesive manner. All the interviews are conducted from different areas of the same BI projects. This enables to get the individual picture that in turn helps to get the overall picture. They can be cross referenced against each other so that they can be analyzed. Hence it allows us to follow the above mentioned methods in chapter 2.4.

All the respondents were approached well before conducting the interviews. They were given short description of the thesis subject area and the research questions for which the thesis is trying to explore answers. All of them know that thesis is going to explore of how to develop BI using Agile Strategies.

The interviews are conducted with open ended questions in-order to gather as much as we can with influencing the respondents. These are the following questions in a broad perspective from the discussions for which we are trying to get answers from the respondents.

1. What do you mean by agile when speaking for BI projects?
2. How do you work with Agile/Lean principles for BI projects?
3. What are important considerations to take into account when dealing with the BI projects in Specific?

After conducting the Interviews they are analyzed and the main points got summarized.
4.4 Information from the various respondent’s

The IT Company that is considered for the empirical study is a very big multinational IT service provider with more than 5500 employees. The Company main focus is to provide service and support to their parent Company which is a product based company. The BI group is a part of the IT Company which provides for the needs of the management BI needs across various companies within the parent company.

Currently during the period of study with the group the BI group members are essentially involved in the Maintenance of existing BI solutions and a project which is going to finish successfully soon. The members of the Project for the first time have successfully used Scrum for the BI project.

The BI group has been using the iterative and incremental development methodology within a gated process framework for the development of their BI projects.

The BI group in the IT organization consisted of Business Analysts, Architects, Developers, Project Managers, Testers and Maintenance managers. The group is mainly involved in supporting the maintenance of the existing BI applications.

I approached them for having semi-structured interviews on various issues of the Agile BI development. When I asked the basic question that If agile BI development possible to implement? The first question that came back was “what do you mean by agile?” Even though it seems that the word agile is popular, it is having a broad meaning. A generic view of agile BI is given which is like this “Agile software development main goal is to deliver the software early, frequent and sustainable delivery of working software that gives business value.”

The major opinion is that they are already implementing agile in developing BI projects as they are using iterative incremental methods for Projects. They also made it clear that they introduce changes to the existing system with no great difficulty hence they are responsive to change. They are currently using user collaboration practices for requirements gathering.

Hence in a way they are already working agile. Everyone feels that agile BI development is necessary and is very useful for developing pro-active systems with in short period of time.

Hence it appears that the development or IT community has already recognized the significance and usefulness of agile strategies and been implementing.

So my next goal was to take a step forward and ask for specific agile methods if they can be implemented. The methods are Scrum and XP practices as these are appointed agile methods in the IT Company under consideration.

OLTP systems and OLAP systems are closely related as both mainly deal with data issues like quality and integration. Hence it’s reasonable to check if their development methods are in anyway similar or if they are in contrast to each other. When I took the same question to a Senior Test he gave the following response.
“I personally don’t see any difference between those two. If there is strong team and working procedure on dealing with the master data management and business team gives us the correct test data based on the Business requirements we can develop easily in an agile way.”

When approached with the same question to a Senior BI Business Analyst he opined in the following way.

“In my opinion nothing, except a more diverse toolset = more division of work amongst team members which is a contradiction to agile methodology

The great difference is that BI requires a thorough understanding of management disciplines which is particularly bad described with common process modeling techniques used in normal software development (transactional processing).
In reality I guess that consultants and other vendor's best practices are also often used”

A senior BI project manager remarked on the differences between traditional BI projects and the BI project that he is managing currently which is using Scrum as follows.

“In traditional BI projects all the activities are done by a small team that is involved in all the phases first the requirements gathering, then the Design, Implementation etc.

But now in scrum we have all the roles divided in between specific people. All working in each sprint and again they involve in other sprints as well. Also we produced needed documentation in each sprint.”

Test driven development is part of the implementing scrum methodology. So when quizzed about the how about of the test driven development a BI Project Manager responded this way.

“We are also doing this Test driven development. At the end of the each sprint the Automated (User Test & Functional Test, regression testing) test cases are attached with the build phase. ”

In brief the project goal for which scrum is being implemented is to attain the Proper visualization tool. And it should provide us with the access to gather the relevant information or data based on the complete Business Specification requirements.

Almost 3 members in the group business analyst, Project manager, Architect felt that the following are critical to the success of the BI project.

“The product owner or the higher authority should commit themselves to the agile way of working.

So they should not freeze the requirements between the gates. Actually that is the way it enables us to be fully agile so that we can have requirements until we reach release gate so that we deliver to the user what we intended to earlier.”
But there is an opinion about the scrum from the whole team.

“The opinion is that the scrum is a great idea.”

A senior business analyst informed that there no BI specific methodology either in the IT company or in the practice. Here is in his own words.

“When it comes to BI development, from my point of view there are not really that many development methods out there on the market. It's really mostly a collection of "best practices" and software and database design patterns.”

When it comes to the tools which support the Agile BI development the following is the response from the project manager which is his experience within the group.

We do not have a tool for the Requirements Traceability Matrix. This helps us to trace and see that all the requirements are unique and their design should be unique and so that they have the unique test cases designed. If we have a tool in the company then it would be great for this.

The main challenge found from the responses from the group is as follows.

“Is that at the end of the each increment we need to incorporate the new functionalities to the application that has not been implemented in the previous sprint.”

A business Analyst from the group has the opinion that BI lacks formal guidelines for working.

“So when comparing "normal" methods with BI, there are no similarities because people tend to work ad hoc in the BI area because of lack of formal guidelines and as far as I am concerned, a lack of understanding of the business needs but it is really a chicken and the egg discussion.”

“But yes, if we have a manager (in this case it can be an application, project or maintenance manager) that takes responsibility of the backlog. These backlog activities need to be aligned and prioritized with all other Sprint activities in the agile way of working. In some cases backlog activities need to be more prioritized than already planned Sprint activities and in that case, the Sprint activities can be postponed to the next sprint instead, or in some cases backlog can be handled within a current sprint.”

The BA gave a feedback that there is a lack of performance measures in development teams at the IT Company.

And also majority of the respondents said that the IT architects and developers should not take back seat position when defining requirements rather should actively participate in requirements defining phase.
A BI BA responded the following on how agile can be implemented

“For what I understand, agile in a broad sense requires pretty much the same techniques as any other project that is correct time estimates, a promise to keep individual commitments and focus on delivery on time but it does not support development in large teams i.e. more than 4 persons so the method does not scale in volume.

The program and project synchronizations needed when creating global enterprise solutions are simply overwhelming. As also the enterprise architecture challenges which takes a lot of thinking and designing to get it right from start? You must understand the scale of complexity we are dealing with here. 200+ systems and hundreds of people involved in projects

The software part does not require any special adoptions in toolsets as far as i am concerned.

We need better support and more discipline to be able to handle configuration and release management tool support and consistent coding. I mean to be able to have effective release build process.”

The respondent is a senior software architect in the field of Data warehousing and BI. He has more than 10 years of experience in this field of BI. He responded the following way when asked for if we can implement a prioritized product backlog.

“Depends on how you mean with “prioritized product backlog”, but I think you mean project backlog – activities that has not yet been developed due to different reasons, time, readiness, complexity etc.

But yes, if we have a manager (in this case it can be an application, project or maintenance manager) that takes responsibility of the backlog. These backlog activities need to be aligned and prioritized with all other Sprint activities in the agile way of working. In some cases backlog activities need to be more prioritized than already planned Sprint activities and in that case, the Sprint activities can be postponed to the next sprint instead, or in some cases backlog can be handled within a current sprint.”

By asking the question if we can implement the Back End functionality (i.e. Dimensional Modeling, Physical design, Data staging design & development) in sprints from 3 to 4 weeks? The following was his response

Yes, most of the cases I think this is possible. Depends on complexity of the delivery, but then you can split it up over more Sprint period with partly deliverables within a 3-4 week period. Most important to be able to work Agile and in sprints it requires that you have an established platform and a solid foundation to build on, I mean an existing BI environment that is up and running, but also a project team that are well aware of the methodologies and common patterns for development with guidelines and check-list on how the development should be done. Else I think it will be hard to deliver in those short sprints.
The business analyst and architect gave a similar responses on how to overcome the agile implementation for the BI.

*Experienced resources are vital that can work together in a team (team-players), this is very crucial. Flexibility in the minds of developers to be able to start working on one task, but need to have the ability to change rapidly if prioritizations are changing due to different facts.*

*Working tightly together, it is more challenging in global teams as you’re not sitting close to each other as communication is very important, not only planned meetings but also the “day-to-day over the office desk communication”.*

*If the existing infrastructure and development tools, guidelines and Patterns for design and development can be re-used then that will speed up the work and enforce faster deliverables in sprints.*

A senior test developer also emphasized the importance of having clear requirements from the business team in a following way.

*“I see that it’s not only specific to BI projects we sometimes deal with these issues in other projects where we need to migrate data from various sources into a single system. We can deal with them in sprints if we have clear requirements from the business team.”*

He also gave his valuable suggestion on how to make the agile way of working in the following paragraph.

*“According to my experience I have worked with the Microsoft BI platform and it’s helped us a lot to be agile. We have good understanding of the product. We have been agile in the development part as we are allowed to be agile from the higher management. In the specified development phase we used Scrum along with Test-driven development and we are successful in that. But to my knowledge we are again the loop having specific roles and requirements freeze before the development phase. Otherwise if the higher management supports and Business community committed to the clear business requirements then it’s easy to be in complete agile way of working.”*

A maintenance assignment leader felt that using Kanban helps a lot to be agile especially for the maintenance tasks of a project. He also said that it reduces double documentation so we can reduce the burden of documentation. He also said *“We need clear & simple policy guidelines to be more agile in very big organizations”.*

### 4.5 Empirical research results

The main purpose of this chapter is to provide answers for questions which will in turn help us answer the main research question.
The main Research Question is

**Q. How do we implement BI development using agile strategies?**

This is answered by answering the following two sub research question.

RQ1. What are the Similarities and differences between the Traditional transactional processing systems development and BI systems development?

*Ans:* There is no fundamental difference between the transactional processing systems and Business intelligence application. Even transactional processing systems projects deals with issues like data quality and data integration where we need to migrate data from various sources into a single system.

In traditional BI projects all the activities are done by a small team that is involved in all the phases first the requirements gathering, then the Design, Implementation etc. But now in scrum we have all the roles divided in between specific people. All working in each sprint and again they involve in other sprints as well. Also we produced needed documentation in each sprint.

There are currently not many development methods available for the development of BI projects in the market. It’s only a collection of best practices. If there is strong team and working procedure on dealing with the master data management and business team gives us the correct test data based on the Business requirements we can develop easily in an agile way. We can even implement the Back End functionality (i.e. Dimensional Modeling, Physical design, Data staging design & development) in sprints from 3 to 4 weeks.

The major difference is that BI requires a thorough understanding of management disciplines which is particularly badly described with common process modeling techniques used in normal software development (transactional processing).

RQ2. How can we adopt to agile way of working for BI development?

*Ans:*

1. There should be commitment and involvement from both the management and business teams.

2. The architects, developers and business analysts should involve in the requirements phase to clearly find out the business needs as it’s our role to find out the exact business needs so that we can deliver an appropriate solution.

3. Test driven development should be done.

4. Requirements should not be freeze down after certain gates which hinder us to be more flexible and agile in order to respond to changes.
5. Face to Face Communication should be done around a desk often. The team should function tightly.

6. The BI platform should already be in place to provide a strong solid foundation for BI.

7. The team should be flexible in mind so that they can welcome unexpected changes.

8. Experienced resources are vital that can work together in a team (team-players), this is very crucial.

9. An experienced project team that has good knowledge about methodologies and common patterns for development with guidelines and check-list on how the development should be done is necessary.

10. Re-use of the Existing Infrastructure and development tools, guidelines and Patterns for design and development can enable us to develop in sprints (i.e. agile).

Tools recommendation:

1. A tool for the Requirements Traceability Matrix is needed.
2. Visualization tool for Kanban will enable us greatly to implement Lean principles.
5 ANALYSIS AND RESULT

In this chapter the Analysis is carried out based on the theoretical and empirical results about the major research questions. The knowledge gap identified in these studies is intended to be filled by the result summary presented at the end of this chapter.

5.1 Analysis

The comparative analysis is used to compare the Theoretical and Empirical Findings with each other. The summary of the theoretical and empirical findings are presented here and compared with each other in the context of answering the main research question. The results of the analysis are presented in the next subsection called result summary.

As mentioned in the section 2.4 above there are several levels of data analysis in the case study which is qualitative. One of the useful actions that will help us a lot is that to arrange the data in a chronological order. This helps us to understand the data and present it in a descriptive manner. It helps to create or develop theories, models and draw inferences.

The results are

The main research question is as follows.

Q. How do we implement BI development using agile strategies?

The answers for the above question will be categorized into 2 sub research questions.

1. Similarities and differences between the Traditional transactional systems development and BI systems development.
2. Combining Agile Methods and BI development methods.

RQ1: Similarities and differences between the Traditional transactional systems development and BI systems development.

From the results of the theoretical study the following similarities and differences are observed from transactional processing systems & BI systems development.

The strategies used for the development of transactional processing systems and BI development systems are quite similar. They have the following phases as common they are Requirements gathering, Design, Development, Implementation and Maintenance.

But when it comes to the application of the agile methods for the development of either OLTP or BI systems the application of agile methods to OLTP systems development is much easier.
The reason being in the OLTP systems development the requirements are often stabilized after certain phases unlike BI systems where there are constant requirements change.

The focus and aim of the BI project differs from that of the OLTP systems. They are as follows.

1. A typical Transactional processing systems goal is to serve the operational data specific needs of the Business. The BI projects goal is to serve the Managers in their decision making process.
2. Orientation towards business opportunities rather than transactional needs.
3. Implementation of strategical decisions, not only departmental or operational decisions
4. Analysis based on business needs, which is the most important of the process
5. Cyclical development process, focused on evaluation and improvement of success

When it comes to the design issues between OLTP and BI systems they are as follows.

1. We must focus on designing for optimum query access. For the operational environment, we also optimize for data manipulation tasks.
2. A data warehouse is developed increment by increment.
3. The final result is a dynamic structure, rather than the static operational structure.
4. The data warehouse stores historical data by time period.

The BI team faces a lot of challenges:

1. It’s unlikely to know what queries are going to be made.
2. Most queries must have high performance with quick response time.
3. Managing the volume of data (incl. Backup and restore).
4. Controlling and managing operating costs.

A BI project is not developing code for the project. It involves lot of Business Data Integration. The Data management is the key important job here that requires nearly 80% of the time consumed by the BI project.

The **Empirical study** provided the following results for the above sub-research question.

1. In traditional transactional processing system development projects all the activities are done by a small team that is involved in all the phases first the requirements gathering, then the Design, Implementation etc.
2. But now in scrum we have all the roles divided in between specific people. All working in each sprint and again they involve in other sprints as well. Also we produced needed documentation in each sprint.
3. There is not much difference among the BI development and classic transactional BI development except a more diverse toolset = more division of work amongst team members.

4. The great difference is that BI requires a thorough understanding of management disciplines which is particularly bad described with common process modeling techniques used in normal software development (transactional processing).

5. So when comparing "normal" methods with BI, there are no similarities because people tend to work ad hoc in the BI area because of lack of formal guidelines, a lack of understanding of the business needs.

6. There doesn’t seem to be any big difference between those two transaction processing system development and BI development. If there is strong team and working procedure on dealing with the master data management and business team gives us the correct test data based on the Business requirements we can develop easily in an agile way.

By using the comparative study for both the theoretical and the empirical research results the theoretical study corroborates the Empirical study. The empirical study makes clear that the transactional development systems and the BI development systems developments are not contrast with each other. They have the same strategies and same phases.

According to the interviewees the BI systems can be developed just like any other transactional systems given they should have more diverse toolset, Formal guidelines, very good management disciplines described with process modeling tools and techniques. The theoretical study opines that the BI projects are more towards business so thus making the application of the agile methods application bit difficult. The Empirical study states this is not big problem because a good understanding of the business by the IT team like Architects and developers enables to develop and deliver a system with agility (i.e. ability to respond to change).

So the theoretical study mainly focuses on the details of how typical transactional systems development differs from the BI systems development. This gap is identified by the empirical study and can be bridged with the capturing of the very good requirements by involving IT team, involving a strong team, a defining a procedure to deal with the master data management, and making sure that the business team giving us the correct test data based on the requirements.

RQ2. Combining Agile Methods and BI development methods

The results from the theoretical study suggested that Agile BI development is about adopting the best practices of agile software development and also with adapting the DW/BI’s own best practices for agility. Agile BI is about early, frequent and sustainable delivery of data in a format that provides valuable business insight.
By using the comparative study on both the theoretical and empirical study the following has been deduced.

The usage of agile strategies for Data warehousing involves with the result of highly iterative approach with collaboration between users, developers and stakeholders.

1. First of all for achieving agility there should be a change in mind set of all the team members. Customer community member’s main intention should be to take sufficient time to explore the newly completed features and give feedback about it.
2. The management community should expect and embrace unexpected change and project risk and uncertainty.
3. The technical team should learn to work in a way which includes lot of discipline and rigor.
4. The project interface community should be ready for daily project involvement and shift in their roles for the overall success of the project.

Important issues that affect the systems for to become a better agile system are as follows.

Collaboration:

1. One of the major reasons identified for the failure of the BI projects is the failure to collaborate; one failure is to collaborate with the customers and the other with the team members.

Evolving Excellent Design:

Evolutionary design involves the following key developer practices. They are

1. Database refactoring
2. Evolutionary data modelling
3. Database Regression testing
4. Configuration Management
5. Developer sandboxes

Evolutionary design involves in producing excellence by effective usage of the agile modelling, Database Refactoring and design patterns. For a system to be able to quickly respond to change and frequently deliver new features requires excellent data models and system design.

Test-driven data warehouse development:

1. For delivering BI systems of high quality in every iteration or sprint requires testing the iteration before delivering. Integration of testing to the development process enables in the delivery of a BI solution of high quality.
2. For a system to be fully agile it is necessary that the routine manual processes to be automated.
3. Hence it is recommended for the DW/BI teams to automate the routine manual processes for continuous integration and deployment and monitoring.
4. Small releases enable the agility of the BI projects.

From theoretical study results it is often identified the Agile BI is a collection of best practices for project management, Design, Architecture. They are actually serving as good foundation on which the empirical results can be structured.

The theoretical results are presented at chapter 3.9 and the empirical results are presented at 4.9. The following table presents an analysis of their results.
<table>
<thead>
<tr>
<th><strong>Empirical Results</strong></th>
<th><strong>Theoretical results</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>There should be a commitment from both the</td>
<td>The theoretical study advocates the strong customer involvement which is a natural outcome from the commitment of the management and the business teams.</td>
</tr>
<tr>
<td>management and the business teams.</td>
<td></td>
</tr>
<tr>
<td>The architects, developers and business analysts</td>
<td>The theory suggests that using User stories can greatly enable the IT team to capture the business requirements needed.</td>
</tr>
<tr>
<td>should actively participate in the requirement</td>
<td></td>
</tr>
<tr>
<td>gathering phase and strive towards understanding</td>
<td></td>
</tr>
<tr>
<td>the business needs of the customer.</td>
<td></td>
</tr>
<tr>
<td>Test driven development should be done.</td>
<td>Theory suggests that Testing throughout the lifecycle and Automate as much as possible. Note: Regression testing should be completed for each change done to the system</td>
</tr>
<tr>
<td>Requirements should not be freeze down after</td>
<td>Theory suggests that Agile BI development is twofold they are Backend implementation (i.e. ETL, DW development) and Front-End development which involves writing code for BI part. The usage of agile even for the backend should be done which makes it completely agile.</td>
</tr>
<tr>
<td>certain gates which hinder us to be more flexible</td>
<td></td>
</tr>
<tr>
<td>and agile in order to respond to changes.</td>
<td></td>
</tr>
<tr>
<td>Face to Face Communication should be done around</td>
<td>Collaboration most preferably face to face communication</td>
</tr>
<tr>
<td>a desk often. The team should function tightly.</td>
<td></td>
</tr>
<tr>
<td>The BI platform should already be in place to</td>
<td></td>
</tr>
<tr>
<td>provide a strong solid foundation for BI.</td>
<td></td>
</tr>
<tr>
<td>The team should be flexible in mind so that they</td>
<td>The theory also supports this finding from the empirical research. It says that Self-organizing and self-managing team will help in emerging best requirements, architecture &amp; design.</td>
</tr>
<tr>
<td>can welcome unexpected changes.</td>
<td></td>
</tr>
<tr>
<td>Experienced resources are vital that can work</td>
<td></td>
</tr>
<tr>
<td>together in a team (team-players), this is very</td>
<td></td>
</tr>
<tr>
<td>crucial.</td>
<td></td>
</tr>
<tr>
<td>An experienced project team that has good</td>
<td>We should have a formal process for deployment and maintenance of the application. This supports the finding from the empirical research.</td>
</tr>
<tr>
<td>knowledge about methodologies and common</td>
<td></td>
</tr>
<tr>
<td>patterns for development with guidelines and</td>
<td></td>
</tr>
<tr>
<td>check-list on how the development should be</td>
<td></td>
</tr>
<tr>
<td>done is necessary.</td>
<td></td>
</tr>
<tr>
<td>Re-use of the Existing Infrastructure and</td>
<td></td>
</tr>
<tr>
<td>development tools, guidelines and Patterns for</td>
<td></td>
</tr>
<tr>
<td>design and development can enable us to develop</td>
<td></td>
</tr>
<tr>
<td>in sprints (i.e. agile).</td>
<td></td>
</tr>
</tbody>
</table>

Table 9 Empirical Findings vs. Theoretical results
By using the comparative study on theoretical and empirical study the following can be deduced. There are a list of Best practices for project management, Design & modeling, Architecture for making the process of BI development more agile.

The primary objective to introduce the agile to DW/BI systems is to deliver frequently a high-quality, high value, working DW/BI system responding to change (Dan, 2010).

As found from the theoretical study that agility is organization specific. Every organization has to define its own agile way of working. Hence there is no a general model or method is proposed for Agile BI development. Hence from the empirical study and theoretical study results the author found that the major practices are around effective user, customer, and developer collaboration. And practices for evolving excellent design and Architecture and Test-driven development are proposed.

5.2 Result Summary

RQ1. What are the Similarities and differences between the Traditional transactional systems development and BI systems development?

*Ans:*

There are not many differences in the traditional transactional systems and the BI systems development.

We use the same strategies for both and they have the same phases of development as any other software project.

If there is

1. More diverse toolset.
2. Good understanding of business needs.
3. Good Formal guidelines
4. strong team and
5. Working procedure on dealing with the master data management.
6. And business team gives us the correct test data based on the Business requirements.

It’s not so difficult for BI systems to be developed much the same way as the traditional transactional development methods which are using agile methods (i.e. iterative & incremental methods).

RQ2. Combining Agile Methods and BI development methods

*Ans:*

The implementation of agile methodology mainly involves bringing in 3 major benefits improved quality, reduced risk and responsiveness to change.

1. For bringing in quality we need to have proper framework testing like Automatic testing, regression test suites etc.
2. Business Involvement and testing throughout the process enables the project not to miss the business expectations. There by quality of the product is ensured and kept high.

3. Frequent delivery can enable to respond to change.

4. In traditional waterfall development the software is delivered in months rather than in months. In agile our goal is to deliver it less than that time.

As it is discussed above that agility is organization specific every organization has to define its own agility by what it means by agile. As mentioned above in the delimitations that the study took a case study into consideration so based on the agile strategies used in the IT Company the following. The Case study is a very big product based company which is using agile strategies for traditional transactional processing systems. The agile strategies they are using are Scrum and XP practices for projects and Kanban for maintenance.

For BI development they are using iterative and incremental methods which have a typical iteration duration which lasts from 90-120 days.

These are the following recommendations done for the making the DW/BI development more agile.

The DW/BI projects initiatives involves with high expectations from the business sponsors as they are critically necessary for today’s competitive business environment.

Often the business sponsors perceive their role to be limited to certain phases of development like initial drive in terms of allocating resources. But the thesis results emphasis that the close collaboration between business users, developers, sponsors is vital. Along with the active participation commitment to the projects from the higher management plays a very important role in the success of the projects.

From the theoretical study for management of the project agile project management is used, which based on the Envision and Explore life cycle. But the IT Company is already implementing successfully scrum methodology for development of the project. Scrum is an agile project management framework.

From the above observations in this study the Envision and Explore life cycle’s main intention is to capture the accurate and precise business requirements. From the theoretical study User stories is recommended for effective capturing of the requirements in the envision and explore life cycle management which can be included within the scrum methodology for the development of the DW/BI projects.

The IT Company considered follows a gated methodology and this methodology restricts the intake of the requirements after certain gates which hinder the project to be more agile. In-order to allow the project to be completely agile the requirements capturing cycle like Explore-envision life cycle should be continued until the end of the project.
A typical BI project can be divided into two parts they are Back-end development or DW development & Front-end development or BI development. The Back-end development is the crucial part which revolves around data integration which consumes 80% of the time and effort.

There are two pre-dominantly used life cycle methodologies Ralph Kimball and Bill-Inmon life cycle methodologies. The theory and empirical study corroborates that ralph Kimball life cycle best suits for agile methods as it involves users for dimensional modeling. For the Back-end to be more agile the re-use of the existing infrastructure and guidelines, patterns and design can greatly enable them to be more agile.

These are the core issues identified upon application of the best practices enables the entire DW/BI application to be more agile.
The core practices presented are tentatively verified.

Figure 8 Core practices for bringing agility to BI development projects.
6 DISCUSSION

6.1 Conclusions

The comparative case study and the Analysis have produced us some interesting facts. They are.

The BI project is typically different from Transactional processing systems hence their development methods also changes in a different way. In a BI project the primary focus is on data. This involves extraction of data from a variety of sources and transforming them and cleaning them and loading them into a data warehouse. And with the high probability of change in requirements during the development of BI project as they support for the management decision making they are developed using iterative and incremental methods.

A typical BI/DW project involves two steps they are 1. Front End development 2. Back End development. In the current the goal of the Thesis is to find out the suitability of the agile strategies specifically Scrum and XP practices for the BI development in the IT Company context. In this theoretical study the author found that the agile strategies are developed mainly by their authors with the focus on code development projects. But the BI projects are often Business Integration Projects. With this in mind it’s often easy deduce from this that the Front End development like Analytic Application development or reporting Application development can be carried out in Sprints(2-3 weeks) and XP(1-2 weeks).

Now the BI practitioners are developing BI projects using the Iterative incremental method which involves iterations of 90-120 days. By using scrum/XP methodologies we can reduce to 3-4 weeks sprints. But reducing the number of days for each delivery of iteration does not make it agile. Agile means collaboration between users, stake holders and developers across cross functional teams. For achieving this we are suggesting the Agile Project Management Framework (i.e. based on Envision and Explore Technique.).

For a system to respond to changes quickly which is one of the attribute to the successful impartation of agility for the project. Evolutionary design involves in producing excellence by effective usage of the agile modelling, Database Refactoring and design patterns.

Having a strong experienced team, and well establish platform and tools enables the team to be more agile but keeping the sprints pretty tight as much of the work can be reused. Test-driven development should be part of the BI development. Along with all the above implementation a strongly dedicated and committed management community is the key to successful implementation of Agile BI projects.

The model presented above gives the list of core practices that should be adopted in-order adopt agility for BI development. The model has not been tested completely. The model was presented to the respondents and they provided with feedback. Hence the model can be adopted tentatively.
6.2 Implications for Informatics

The case study is conducted at a very large multi-national IT company. The data acquired by the empirical research is from this IT Company. So the data is pretty much relevant to the field of informatics as the data origination is from informatics area.

The field of informatics main focus is the transformation of information either by Computation or Communication whether by Organisms or Artifact’s. The result presents the effective means of developing the systems which transform the raw data into information and there by presenting it as knowledge to managers, researchers, business people etc. who are in need of that information.

The result elicits the many benefits of these systems development using agile as they reduces waste, and adds agility or responsiveness to the system which alleviates the user profitability.

With the growing size of the data in the form of structured and unstructured data and with ever growing competition knowledge becomes indispensable for informed decision making. The demand for the development of the systems (BI) which turns data to information and information to knowledge is growing. The systems should be able to respond to change as they need assist in decision making. The whole concept of agile BI is to build BI systems with more responsiveness to changes within short delivery cycles. Hence the results directly contribute to field of informatics.

6.3 Method evaluation

The theoretical study’s main focus is to find the relevant subject areas and connect it to the research questions and see how they answer them. This was done by digging into the previous research on the topics and sees how we can build the case of Agile BI effectively. The journal articles, books provided to be a secondary data for answering the research questions.

The theoretical results were verified by the interviews and observations conducted within a BI team in a very big IT company. The theoretical findings served the purpose for which it was conducted. But still there are some areas where theoretical results alone cannot present the complete picture or answers. The observations even though concede to that of the theoretical findings but miss the mark on how to pragmatically implement them. This gap is filled by the Empirical findings.

The empirical findings main purpose is to answer the research questions. Hence the findings are structured in way to so as the various responses for the various respondents confer the research questions. Hence the perspective from the empirical findings and theoretical findings are compared to each other using comparative analysis. Hence this gives us the picture on how to adopt agile way of working as the theory and empirical findings complement each other so as to provide us the complete picture.
6.4 Result evaluation

The research aim is to find how agile can be adopted for BI development. The data analysis has been used to search articles from various journal publications.

Triangulation method is used to compare data from various sources. The data from the primary and secondary sources are compared. The patterns and themes are drawn from the comparisons and results are presented. There are also four criteria that are used for the evaluation of the theory and empirical results. They are triangulation method, external reliability, internal reliability, internal Validity.

Internal Reliability:

This deals with the consistency of the data collection, analysis and interpretation. The theory is built on the scholarly articles from various specialized journals from the field of information systems. And also the articles are also filtered from an online library articles filtering tool called as summon. The articles are chosen using their relevance to answering to the main research question. As mentioned in the 2nd chapter the data from the theoretical and empirical findings is arranged in a chronological order so as to present the data in a descriptive manner. The respondents agreed most of the times during their responses. Then the analysis and interpretations are done. Hence the Internal reliability for this research is valid.

External Reliability:

This is about “If the Extent to which independent researchers can reproduce a Study and obtain results similar to those obtained in the original study.” The research provides with how to adopt the agile way of working for BI. Hence if the theoretical and empirical findings are compared in the context of the specific organization there is a greater chance to respond in the same way as this research is on bridge the gap in implementation

Internal Validity:

The internal validity is valid for this research. The observations from BI group in the IT Company and the theoretical results agreed to major extent. Hence the Internal validity is also high.

6.5 Possibilities to generalize

This is about the generalization of the result outside of the selected sample. The main research focus is on agile BI development. The theory suggests gives us the list of Best practices in project management and Architecture and tools for the BI development. But the Empirical findings provided us with the facts in the real time implementation and how can it be practically implemented and how did they felt during the development phases. The theoretical findings and the empirical findings provide very good solid ground for the BI development. The observation found by working closing to the group enabled to provide the real time reflections on what are the differences in Traditional OLTP systems development and the OLAP systems development. And also this served to bridge the gap on how to adopt the agile way of working in BI. Hence even though the agile methods and development tools and
platform are specific to the Company under sample the results can be generalized in a broader context.

The goal for implementing Agile BI solution development is same everywhere hence the results can be used when applying agile for any specific organization. This enables the development team and management to consider issues that are important when going for agile way of working.

6.6 Ideas for continued research

The study focused on how to implement the agile method for BI development. The research considered scrum as agile method and some XP practices. The organization under sample is currently only using scrum and XP practices. As part of the work this also found that how to implement these specific method of weekly sprints to BI development. But the continued research on how to adopt different agile methods to BI development could be very much useful.

The research proposed a model of core practices in order to be implemented for Agile BI development. That can be verified across various organizational contexts and adopt accordingly with necessary changes.

The research listed a core practice named “Test-driven development”. This is a very interesting and useful area which can be studied further to present in detail on the efficient methods and best tools available for implementing this.

The core practice named “Excellent design” is a very important area to be studied further on the challenges of developing an excellent design and suggest necessary tools and methods.

The mindset of the people in the project impacts a great on the adoption of agile. The topic on how to change the mindset of the project team to suit agile can be an interesting area to investigate.

6.7 Speculations for the future (if any)

The research has given provided various means on how to effectively implement agile strategies in BI development. There is a strong support from theoretical and empirical findings that the concept of agile BI is very important. As the concept of agile mean different for different companies and development team there should always their own way of implementing for their purpose. In this research the company considered as sample considered scrum, XP practices as agile on Microsoft platform and they are successful in implementing them for agile. But when implementing other agile strategies in various other platforms there is a chance for the results to vary. Hence it’s important to investigate the impact of agile strategies and their specific platforms and adopt it their specific organization culture. So this research work should be updated with latest agile methods and their implications in various organizations who are implementing BI.
7 References


Introna, L. D. Information: A hermeneutic perspective.


Merriam, S. B. (2002). Introduction to qualitative research. Qualitative research in practice: Examples for discussion and analysis, 3-17.


Strode, D. E., Huff, S. L., Hope, B., & Link, S. (2012). Coordination in co-located agile software development projects. 85(6), 1222-1238. Retrieved from http://hb.summon.serialssolutions.com/link/0/eLvHCXMcwVV05DgJBDJxPSUELNB0aKMjvH1ogVD4AH5Jqsiv8LL6KAD6RyYltv5JTOoBIrRh3QbQ6NRt2jhDp1gtRB5_Wtu-Lnm2zE9tu9csvfHoAccFCceVZlFZldCtQ2s_Vo-9CidTVa4MPLHOFAL9jGlz6UTNaYhk1XYPaUDrLnxZ-vz1-ZvwFJYCIO. doi:10.1016/j.jss.2012.02.017


University of Borås is a modern university in the city center. We give courses in business administration and informatics, library and information science, fashion and textiles, behavioral sciences and teacher education, engineering and health sciences.

In the School of Business and IT (HIT), we have focused on the students' future needs. Therefore we have created programs in which employability is a key word. Subject integration and contextualization are other important concepts. The department has closeness, both between students and teachers as well as between industry and education.

Our courses in business administration give students the opportunity to learn more about different businesses and governments and how governance and organization of these activities take place. They may also learn about society development and organizations' adaptation to the outside world. They have the opportunity to improve their ability to analyze, develop and control activities, whether they want to engage in auditing, management or marketing.

Among our IT courses, there's always something for those who want to design the future of IT-based communications, analyze the needs and demands on organizations' information to design their content structures, integrating IT and business development, developing their ability to analyze and design business processes or focus on programming and development of good use of IT in enterprises and organizations.

The research in the school is well recognized and oriented towards professionalism as well as design and development. The overall research profile is Business-IT-Services which combine knowledge and skills in informatics as well as in business administration. The research is profession-oriented, which is reflected in the research, in many cases conducted on action research-based grounds, with businesses and government organizations at local, national and international arenas. The research design and professional orientation is manifested also in Innovation Lab, which is the department's and university's unit for research-supporting system development.