Data warehouse development
- An opportunity for business process improvement

Jesper Holgersson

Department of Computer Science
University of Skövde, Box 408
S-541 28 Skövde, SWEDEN

HS-IDA-MD-02-006
Data warehouse development
- An opportunity for business process improvement

Submitted by Jesper Holgersson to the University of Skövde as a dissertation towards the degree of M.Sc. by examination in the Department of Computer Science.

October 2002

I certify that all material in this dissertation which is not my own work has been identified and that no material is included for which a degree has previously been conferred on me.

Signed: ________________________________
Abstract

Many of today’s organizations are striving to find ways to make faster and better decisions about their business. One way to achieve this is to develop a data warehouse, offering novel features such as data mining and ad hoc querying on data collected and integrated from many of the computerized systems used in the organization. A data warehouse is of vital interest for decision makers and may reduce uncertainty in decision making. The relationship between data warehousing and business processes may be used at the pre-deployment stage of a data warehouse project, i.e. during the actual development of the data warehouse, as an opportunity to change business processes in an organization. This may then result in improved business processes that in turn may result in a better performing data warehouse. By focusing on the pre-deployment stage instead of the post-deployment stage, we believe that the costs for development will decrease, since needs for changes detected early in a development project probably will be detected anyway, but in a later stage where changes in the business processes may cause a need to restructure the finished data warehouse. We are therefore interested in which factors that may cause a need for changes in the business processes during the pre-deployment stage of a data warehouse project, the types of business processes affected, and also if there is any correspondence between factors that trigger changes and business processes affected.

Based on a literature survey and an interview study, general triggering factors to change business processes have been identified, such as needs for new organizational knowledge and for prioritization of goals etc. We have also found that needs for changes more often concern supporting processes than other types of business processes. We have also found a general correspondence at a type level between triggering factors and affected business processes.

In combination with the results and conclusions presented, we have also identified propositions for future work, which will refine and confirm the ideas presented here.

**Keywords:** Data warehouse, Business process, Business change.
Acknowledgements

There are several people I would like to thank for their support during hard times of work of this M.Sc. year. First of all, I thank my supervisor Mattias Strand for his valuable comments, guidance and engagement in my work. I also want to thank Anne Persson for valuable comments. Furthermore I would like to thank Lars Boström for the help with reaching respondents for my investigation. Finally, I would like to thank my family for their support. Without all of you, this work would not have been possible.
# Table of contents

1 Introduction ........................................................................................................................... 1  
  1.1 Problem area .................................................................................................................. 1  
  1.2 Aim and objectives ........................................................................................................ 2  
  1.3 Research approach ........................................................................................................ 3  
    1.3.1 The research process .............................................................................................. 3  
    1.3.2 Interviews ............................................................................................................... 4  
  1.4 Main contributions ........................................................................................................ 5  
  1.5 Thesis outline ................................................................................................................. 6  

2 The data warehouse .............................................................................................................. 8  
  2.1 What is a data warehouse? ............................................................................................ 8  
    2.1.1 Characteristics of data warehouse data ................................................................. 9  
    2.1.2 The difference between a data warehouse and OLTPs ...................................... 10  
    2.1.3 Benefits of data warehousing .............................................................................. 12  
    2.1.4 Problems of data warehousing ............................................................................ 13  
  2.2 Developing a data warehouse ..................................................................................... 14  
    2.2.1 The project planning phase .................................................................................. 15  
    2.2.2 Defining the business requirements .................................................................... 16  
    2.2.3 The technology track ........................................................................................... 17  
    2.2.4 The data track ....................................................................................................... 18  
    2.2.5 The application track ........................................................................................... 19  
    2.2.6 Project management ............................................................................................. 20  
    2.2.7 Deployment .......................................................................................................... 20  
    2.2.8 Maintenance and growth ..................................................................................... 21  
  2.3 Data warehouse architecture ....................................................................................... 21  
    2.3.1 Operations and tools in the data warehouse ....................................................... 22  
    2.3.2 OLAP servers ....................................................................................................... 23  
    2.3.3 Data marts ............................................................................................................. 24  
    2.3.4 Front end tools ...................................................................................................... 25  

3 Processes ............................................................................................................................. 27  
  3.1 Business processes ....................................................................................................... 28
1 Introduction

In today’s organizations, there exist numerous kinds of computerized systems that keep track of and manage different kinds of data. One such type of systems is data warehouses. Chaudhuri & Dayal (1996) describe a data warehouse as a database system that integrates data from various sources. The usage of data warehouses has previously mainly focused on getting knowledge about customers etc. in order to reach new market shares and thereby selling more products. Nowadays, the usage of a data warehouse has been broadened and in house organizational usage of the data warehouse has been added on to the usage areas. According to Agosta (2000), a data warehouse system (DWs) supports decision making in different types of business processes, which indicates that there is a strong relationship between data warehousing and business processes. However, data warehousing is, as stated earlier, not only aimed at supporting the decision making in different business processes. Data warehouses and their development projects may also be used when redesigning or changing business processes. Stated in another way, i.e. Kelly’s way (Kelly, 1996, p. 66):

“[…] the data warehouse is becoming a key driver of process redesign in the business, and this is likely to be a common feature of data warehouse implementations.”

1.1 Problem area

From the above statement the conclusion may be drawn that the relationship between data warehousing and business processes can be used as an opportunity to change or redesign business processes in an organization. The outcome may be better designed business processes, which also may increase the performance of the data warehouse. Another obvious outcome may be that better designed business processes make the execution of operations in the organization more efficient.

Redesign of business processes may occur at two stages during a data warehouse project. It is worth mentioning that this partitioning of data warehouse projects is very rough, but the partitioning given is useful to show the two main parts of a data warehouse project that are of interest in this thesis.
1 Introduction

- The pre-deployment stage of a data warehouse project, which concerns the analysis of the organization and its business processes. In this stage, it is possible to detect a need for data that do not exist in the organization today and that has to be derived in order to satisfy the requirements of the data warehouse.

- The post-deployment stage of a data warehouse project is, according to Kelly (1996), a stage that concerns the actual usage of the already implemented data warehouse. The data warehouse may identify new potential strategies for the organization and these new strategies may cause a need for redesign of the organization’s processes.

Agosta (2000) claims that the development and maintenance of a data warehouse costs five times more than the hardware and the software. Therefore we find it interesting to investigate, if we in the early phases of a data warehouse project (pre-deployment stage) can detect a need to change business processes, that otherwise would be detected in the post-deployment stage of the data warehouse project and thereby cause increased development costs since changes at this stage automatically will affect other parts of the project. Another interesting aspect of the subject is that Connolly et al. (1999) mention that there are some problems associated with the development of a data warehouse. Connolly et al. (1999) claim that during the development of a data warehouse, it is possible that there may emerge some previously unknown problems in the source systems, such as business processes that cannot produce the required data. We argue that there is a possibility that an early discovery of these business processes will decrease the costs and efforts for developing a data warehouse.

1.2 Aim and objectives

To address the issues mentioned above the work reported in this thesis aims to:

*Investigate the possibility to identify which factors may cause a need for changes in the business processes during the pre-deployment stage of a data warehouse project.*
1 Introduction

In order to fulfill this aim, the following objectives must be achieved:

- Identification of particular types of business processes that are affected by the introduction of a data warehouse and also how these business processes are affected.
- Identification and categorization of the factors that cause the need for business process change during the pre-deployment stage of a data warehouse project.
- Examine if there is any correspondence (type of business process affected by a certain factor) between factors that cause changes, and the types of processes that are affected.

1.3 Research approach

A work process was designed to give the intended work a solid structure, and to minimize the risk that some aspect of the aim and objectives would remain uninvestigated. The work process also gave details for how the interview part of the research project would be conducted.

1.3.1 The research process

In order to carry out the intended work, we formulated a work process. This process outlined how the actual work should be conducted. The order of activities at the work process was not considered as static. Instead, in order to be more efficient, some of the activities outlined were conducted in parallel. The working process is presented graphically in figure 1.

Firstly, a literature study was done. The literature study aimed at laying the foundation for further investigation. It also provided a good overview of data warehousing and business processes, and clarified concepts and ideas within these areas.

Secondly, interviews were conducted to collect material needed for being able to answer the objectives stated in section 1.2. We have chosen interviews since we were interested in relatively detailed information from different data warehouse development projects.
Finally, another literature study was performed. This was used when the collected material was analyzed. By doing this, more information about collected material was found.

Figure 1: Work process

1.3.2 Interviews
The purpose with the interviews was to collect information related to the aim and objectives. This information was intended to work as an input to the second literature survey. To achieve this information, we used a flexible questionnaire, which allowed a discussion of certain details if this was considered necessary.

The questions used in the interview had a low degree of structure. A low degree of structure gives, according to Patel & Davidson (1994), freedom when answering questions, i.e. the respondents will not be restrained when answering questions.

Interviews can be made in different ways. One such way is face – to – face interviews. Dahmström (1991) mentions some benefits and shortcomings associated with this way of
making interviews. A fundamental benefit with face – to – face interviews is the possibility to have a deeper discussion of the subject and thereby reach information that would be hard to get in another way, such as questionnaires. A shortcoming with face – to – face interviews is that the material may be harder to analyze because of its great extent and will, in combination with traveling, require more resources, i.e. time and money.

Another way of making interviews is to use telephone. An interview made by phone contact between respondent and interviewer will save a lot of resources. A shortcoming is that it is hard to get a deeper discussion about the subject, and thereby some information may never be found.

Since we soon found out that the respondents did not have the time required to participate in face – to – face interviews, we made the choice to go for telephone interviews. We do not believe that this choice makes the interview study less appropriate, since the questions asked were highly standardized and had a low degree of structure.

We thought there was a chance that different projects would have quite different information to offer, since the data warehouse development projects are specific for a certain organization. For this reason we used respondents that had participated in a number of different projects in different organizations, in order to get a more general view.

The respondents that were used were found by using a contact person, who is considered to be a well known actor in the area of data warehouse development. This contact person involved well experienced actors, with respect to data warehouse development. The interviews were held in Swedish and recorded. The recorded interviews where then translated and transcribed to English. All statements from the respondents cited in the thesis have been translated into English.

1.4 Main contributions
The main contribution from this thesis is the highlighting of the importance to pay attention to the change of business processes early in the development of a data warehouse. We
1 Introduction

argue that this awareness will reduce the complex environment of the development and thereby make the building of data warehouses easier and less costly.

Another contribution made in this thesis is a characterization of factors that may affect business processes during the pre-deployment stage of a data warehouse project. We have shown that triggering factors at a general level apply to all types of business processes and have a lot in common with factors that cause change in other types of Information Systems Engineering (ISE). Examples of general triggering factors are a will to reach maximum return on investments, prioritizing of goals, new organizational knowledge etc. At a type level, we have derived some correspondences between triggering factors and affected business process types. We have also concluded that there is no correspondence between triggering factors and affected business processes at an instance level, since these are considered to be specific for certain organizations and projects.

Types of business processes affected by a need for change in the early phases of a data warehouse project makes clear that all types experience a need for change. Although, we have made clear that organizations handle this need differently, depending of which type of business process that is considered. Core processes are considered harder to change, compared to supporting and management processes, because of their central role in the organization.

We also believe that this thesis will make a contribution by paying more attention to the business processes in an organization, when developing a data warehouse and give a new way of viewing the importance and treatment of business processes during the development of a data warehouse.

1.5 Thesis outline

In chapter 2, we give an introduction to data warehouses and discuss definitions, characteristics, and architecture. The architecture is explained in general, without any deeper details, since the aim for this work does not concern the actual behavior of a data warehouse. Still, we consider it to be important to give the reader some information of the concepts involved in data warehousing. We also give a brief introduction of how to develop
1 Introduction

a data warehouse. We have chosen one way of performing a data warehouse project, which is commonly mentioned in the literature. This widespread view is complemented with other author’s views of data warehouse projects. This will ensure that an objective and well accepted description of a data warehouse project is presented.

Chapter 3 discusses the concepts of business processes and states what distinguishing features that represents business processes. In this chapter we also point out different kinds of business processes that have been identified in literature related to the topic.

Chapter 4 concerns how the information, which is used in order to solve the aim and objectives of this thesis, has been collected. We discuss the type of questions used during the interviews and associated to this; we make a proposition to how the collected information may be analyzed.

Chapter 5 presents the material collected during the interviews. Firstly, the respondents are presented and their experience of data warehousing, views of business processes and data warehousing is discussed and illustrated with statistics. This demonstrates that selected respondents are suitable for participating in the investigation. We then discuss the material collected that is related to each of the objectives given in this thesis, i.e. types of business processes affected, triggering factors and the correspondence between triggering factors and business processes.

The analysis (chapter 6) is discussing the information described in chapter 5. The analysis contains arguments and explanations to information found from interviews. The information analyzed, is also complemented and contra stated with ideas from related literature, which gives the analysis a more solid ground. Chapter 7 presents results derived from previous chapters. Chapter 8 discusses the work that has been done in this thesis. Different problems are highlighted and the main contributions are discussed in order to place the new ideas presented in this thesis in its right context. Chapter 8 also includes suggestions for future work, which will take the ideas presented in this thesis further and probably add new knowledge that has not been revealed in this thesis.
2 The data warehouse

Chaudhuri & Dayal (1996) mentions that organizations have struggled to develop computer systems that automate business processes. Companies that have been successful in their development of computerized systems have gained competitive advantage. Inmon et al. (1999) state that these early computer systems were perfectly suited for gathering and storing data, but they were not adjusted to analyze the data. Nowadays companies are searching for an approach to use operational data to support decision making, which is supposed to result in new competitive advantage (Harding & Yu, 1999). However, since this work is focusing on the problem of deploying a data warehouse at a general level, there is no need for a detailed account of data warehousing. Although, we intend to give a brief high level description of technical aspects, since we believe that this will give the reader a better overall picture of the work.

2.1 What is a data warehouse?

There exists numerous definitions of what a data warehouse really is and this implies that it is not easy to clearly define the concept. Kelly (1996) defines a data warehouse as a single integrated store of data. This store provides the infrastructure for informational software applications in an organization. Chaudhuri & Dayal (1996) define data warehouse as a collection of decision support technologies, which are aimed to support knowledge workers when making decisions. Inmon (1993) means that a data warehouse shall support management in their decision making process. Kimball et al. (1998) define a data warehouse as the query able source of data in the company. Another interesting definition is presented by Strand (2000, p. 11):

“A data warehouse is a set of databases with supporting tools to collect and manage data and to provide strategic and tactic levels decision makers with clean, consistent and aggregated data from multiple remote and heterogeneous information sources.”
2 The data warehouse

Strand’s (2000) definition implies that mainly middle or upper level management in an organization uses a data warehouse and that data found in the data warehouse is integrated from several sources.

On the contrary, Connolly et al. (1999) mention that different definitions of data warehouse are irrelevant, since the goal with a data warehouse always is the same. That goal is to integrate selected data into one single warehouse. The user may then access this single storage in order to pose queries, produce reports and perform analyzes. Inmon et al. (1999) support this statement and claims:

“The data warehouse unlocks the data jailed in the corporate bowels and allows easy and unfettered access”

(Inmon et al. 1999, p. xiii)

The definitions presented concern the actual use of a data warehouse, after it has been implemented. We consider the definitions given to be appropriate in order to describe what a data warehouse is and what it is used for. As a complement to the definitions given, we find it necessary to describe, in brief manners, the technologies and the terminology about a data warehouse. This is done to give the reader some knowledge about how a data warehouse really works.

2.1.1 Characteristics of data warehouse data

According to Breitner (1997) a data warehouse is the connection between organizations operational data store and decision support systems. The data in data warehouses have some unique properties compared to data in ordinary operative databases. In accordance with Inmon (1993), the data in a data warehouse is subject – oriented, integrated, time – variant, and non – volatile.

- Subject – oriented means that the data is structured around the core subjects of the company, i.e. customers and products, instead of the major application areas that are
The data warehouse common in On Line Transactional Processing systems (OLTPs), such as ordering systems.

- Integrated represents the convergence of data from different computerized systems. The data that is being integrated is often inconsistent in different ways, such as different field lengths, and it is important that the integrated data source is being made consistent, otherwise the users of the data warehouse are provided with non-unified views of the data in the data warehouse, which makes the data hard to access.

- Time – variant means that data is connected to a certain time period. This means that data only is valid at a certain time point, or over some time interval. Another important feature of data in data warehouse is that it is stored for a long time. Gray (1999) claims that data in a data warehouse is stored for up to ten years, compared to 60 – 90 days for data in OLTPs. This implies that it is possible to pose queries that concern a specific time interval in a data warehouse. The long storage time results in that the data is historical and this gives, according to Gray (1999), a good opportunity to analyze trends.

- Non – volatile means that once that data has been stored in the data warehouse, it cannot be changed. Although it is possible to update the data and integrate it with the existing data. As a result of this, access to data in a data warehouse is executed in a read – only mode.

The above characteristics seem to be generally accepted and have been used by Chaudhuri & Dayal (1996) and Kelly (1996). However, Kelly (1996) further mentions another feature besides those mentioned above, that is obvious but still worth mentioning. The data in a data warehouse is separated from the operational systems in the company. The data from the operational systems in the company is a by – product and is not used after the transaction has committed. It is this data that is stored and reused in the data warehouse.

2.1.2 The difference between a data warehouse and OLTPs.

The data in a data warehouse consists of data from various data sources, such as OLTPs. These systems are used in the day-to-day operations of the company. Agosta (2000) claims
that OLTPs are transaction systems that perform daily operations in the organization, such as business events related to customers. Examples of source systems are transaction systems, ERP systems (concerns different flows in an organization) and E-trading systems (trading systems used on the Internet). Connolly et al. (1999) give the following description about differences between OLTPs and DWs, which clearly identifies the different properties between the two systems.

Table 1: Differences between OLTPs and DWs (Connolly et al., 1999, p. 916).

<table>
<thead>
<tr>
<th>OLTP – systems</th>
<th>Data warehouse – systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stores present data</td>
<td>Stores historical data</td>
</tr>
<tr>
<td>Data is dynamic</td>
<td>Data is largely static</td>
</tr>
<tr>
<td>Application – oriented</td>
<td>Subject oriented</td>
</tr>
<tr>
<td>Supports day – to – day Decisions</td>
<td>Supports strategic decisions</td>
</tr>
<tr>
<td>Serves a large number of operational users</td>
<td>Supports a small number of managerial users</td>
</tr>
</tbody>
</table>

Table 1 clearly differentiates between OLTPs and DWs. OLTPs stores data that is used on a day – to – day basis of organizations, and the data is always up to date. The DWs on the other hand, stores different versions of data that has been used by the OLTPs. The data in OLTPs can be considered dynamic, since users pose queries that are rather simple and non-analytic. The data warehouse performs another type of queries that are more complicated and analytic, which results in that data must be of a more static nature. OLTPs are application oriented, in the sense that they store individual information that is dispersed in the system. The DWs is instead subject oriented, which means that the DWs focus on the main areas in the organization. For instance may a DWs focus on areas such as customer and products. An application oriented approach would instead focus on individual information, such as prizes and telephone numbers. OLTPs are said to support day – to – day decisions and this means that the users of the OLTPs applies the stored information in their daily work. An example of day – to – day decisions may be that a person at production wants to know the present number of a certain item. DWs are used in a different way, since
they support strategic decisions. The strategic decisions do not concern the present number of stored items. Instead, historical data is used to do decision making easier and more reliable. Finally, OLTPs are used by a large number of operational users that uses the OLTPs in their daily work. The DWs is used less frequent by a smaller number of users that are responsible for the decision making in the organization.

2.1.3 Benefits of data warehousing

According to Kelly (1996), new needs in companies have aroused. Simply satisfying the customer is no longer acceptable. Instead it is necessary to delight the customer. Keeping up with the competition is no longer a guarantee for survival. What is important is to surprise the competition. In a mass customized market, analysis of data patterns will become a revolutionary tool. Breitner (1997) implies that a company’s ability to compete is determined by the management’s ability to deliver exact and fast decisions, and these decisions must be based on information that is complete, of a high quality and up to date. Kelly (1996) mentions three main reasons for introducing a data warehouse: pressure from competitors, mass consuming markets that want to be more customer – oriented, and finally, change of the internal processes in the company.

Connolly et al. (1999) mention some major benefits from a successful implementation of a data warehouse:

- Good returns on investments of a data warehouse. The cost for implementing a data warehouse in a company may be huge (£50,000 - £10 million, according to Connolly et al. (1999). However, the returns on these investments (ROI) are in most cases larger and worth the investments.

- Competitive advantage for companies that has successfully implemented a data warehouse is another important feature. The competitive advantage stems from that new, previously unknown, or previously unavailable information is detected as a result of introducing a data warehouse, such as new information about customers and previously unknown customer trends.
2 The data warehouse

- Increased productivity of corporate decision makers means that decision makers are presented with an integrated unified view of the company’s data systems and this makes decision making more precise and efficient.

Agosta (2000) complements the above statements by pointing out that a data warehouse offers a technological framework that supports the fulfillment of business goals in a cost effective manner and short time. From this statement it can be said that the introduction of a data warehouse gives a company abilities to make faster and better decisions and can therefore increase their opportunities to compete on the market.

2.1.4 Problems of data warehousing

Connolly et al. (1999) also mention some problems associated with developing and managing a data warehouse. The most interesting problems mentioned with respect to the aim of this work, will be further examined:

- Underestimation of resources for data loading is one problem that affects the actual development of a data warehouse. A frequent mistake made by developers is that the time required for extraction, cleaning and loading the data into the data warehouse is miscalculated. According to Inmon (1993), this process may account for 80% of the development time, and it is important that an organization developing a data warehouse is aware of this.

- Another problem associated with the development of a data warehouse is that it may reveal some previously unknown problems in the source systems, such as for instance badly designed business processes that cannot deliver the required data. It is possible that an early observation and correction of these badly designed business processes will decrease the costs for development and maintenance.

- Associated with previously mentioned problems, there may be a problem that required data is not captured by the source systems. This problem forces the organization to make a choice. One way to go is to modify the source systems, and the other way to go is to create a new system that captures the missing information.
2 The data warehouse

In both cases, the development of the data warehouse will slow down and cause increased costs for the organization.

In summary, the problems mentioned above reduce or even prevent that the development of a DWs is successful. It is important that organizations aiming for the development of a data warehouse are aware of the problems, and strive after decreasing these problems within their development project.

2.2 Developing a data warehouse

According to Kelly (1996), it is necessary to define the information requirements of an organization as a whole, before the actual data warehouse development project kicks off. One way of doing this is according to Kelly (1996), to characterize the business at an enterprise level in the context of information characteristics. This reasoning is also supported by Agosta (2000). However, Kimball et al. (1998) pay more attention directly to the business requirement definition. Still, there is a strong consensus amongst the authors that the development of a data warehouse begins before the physical development of the data warehouse takes place. Kimball et al. (1998, p. 33) illustrate the development of a data warehouse in the business dimensional lifecycle diagram:

![Figure 2: The business dimension lifecycle diagram](image-url)
The lifecycle diagram covers all the phases that need to be fulfilled when developing a data warehouse. One important aspect of the business dimension lifecycle diagram is that it contains three parallel tracks, which are called data track, technology track and application track. These tracks are complemented with project management. Kimball et al. (1998) pay attention to that the given tracks have to be developed concurrently in order to achieve a good development process of a data warehouse. We intend to briefly explain every track and their related phases of the business dimension lifecycle diagram in order to introduce the concepts and terminology during the development of a data warehouse. We consider this to be an important aspect since the aim for this work is concentrating on affected business processes when developing a data warehouse.

The business dimension lifecycle have a lot in common with any ISE methods and includes organizational analysis, requirements specification etc. The main difference between ordinary ISE methods and the business dimension lifecycle is the larger focus on data, which is considered to be more important in data warehouse development than in other types if ISE.

Several authors have proposed different ways of developing a data warehouse. We have chosen to focus on the description made by Kimball et al. (1998), since this description contains both the pre-deployment and post-deployment stage mentioned by Kelly (1996), but describes mainly the pre-deployment stage in more detail, which is considered as a contributory, since Kelly’s (1996) description of a data warehouse project is rather roughly partitioned. In order to present a unified view of the development of a data warehouse, we have chosen to add other related literature in the area to the business dimension lifecycle presented by Kimball et al. (1998). In order to avoid mentioning the reference to Kimball et al. (1998) during the whole chapter, we only point out other references. All the other material is collected from Kimball et al. (1998).

2.2.1 The project planning phase

The first phase in the business dimensional lifecycle is the project planning phase. The project planning phase addresses the definition and scoping of the data warehouse project,
and this also includes the motivation and justification of the data warehouse. Glassey (1998) stresses the importance of justifying the data warehouse with respect to the users. Glassey (1998) continues by stating that the underlying architecture and technology of a data warehouse may be ever so good, but if the users feel resistance to the system, it will not be used and thereby the data warehouse will not be a profitable investment.

With the exception of addressing the justification of the data warehouse, the project planning phase focuses on resource and skill-level staffing requirements on each of the phases in the business dimensional lifecycle. This will serve as a guarantee and support for the whole development project of the data warehouse.

2.2.2 Defining the business requirements

According to Glassey (1998), the probability for a successful data warehouse is increased by getting a good understanding of the business and its end-users. Without this knowledge, it is likely that the development of the data warehouse will be an exercise in futility for the developers. The way to collect information from knowledge workers is very different compared to a more data driven requirement analyze of an organization. It is important to stress that the data warehouse designers must understand the key factors that conduct the organization forward. By understanding this, it will be easier to determine the business needs and translate these needs into the data warehouse design. Agosta (2000) have the same basic ideas about how to define the business requirements, but states more explicitly that it is very important that the users of the system, in an early phase of the data warehouse project, get a chance to formulate their requirements of the system. The defined business requirements build the foundation for the following phases of the business dimensional lifecycle.

To summarize, the earlier phases are crucial since it is important that the developers get a good overall picture of the organization and its core business processes. By doing this, it will be easier to make a good alignment between the business and the data warehouse. Another important aspect in the earlier phases is that the whole data warehouse project is planned carefully. The effect may otherwise be that required resources are not planned for,
which in turn will slow down the development project. Moreover, it is also important that the business requirements are stated. Otherwise, the developers will not get a good understanding of the business, and thereby it will be hard to translate the business needs into the data warehouse design.

2.2.3 The technology track

The technology track consists of two phases: technical architecture design and product selection and installation. In technical architecture design, the overall architectural framework and vision is established for the data warehouse. This means that the actual physical design of the data warehouse components are stated and the overall function and aim for the data warehouse is settled. There are three factors that need to be highlighted during the technical architecture design; the business requirements, the current technical environment and planned strategic directions. The business requirements contains, as stated above, a good understanding of the business and the end – users. The current technical environment is also important since it is necessary to have a good picture of how the different source systems work, since they are the sources to the data warehouse. When establishing the architectural design for the data warehouse, it is critical to know how the current systems in the organization are working. The last factor is planned strategic decisions, which is important since the developers of the architecture and vision for the data warehouse need to be sure in what way the data warehouse is intended to be used and what type of support for the organization the data warehouse is supposed to offer. These three factors must be considered simultaneously if the technical architecture design shall be successful.

The second phase in the technology track is product selection and installation. In the previous phase, the technical architecture design was specified. The product and installation phase uses the technical architecture design as a framework for selecting specific components in the architecture such as database management system (DBMS), hardware platform, query engines etc. The developers shall evaluate possible alternatives of specific hardware components in order to find out which component that best fits the needs for the project. Once the appropriate components have been selected, they are installed and tested
to make sure that their integration with other components in the data warehouse is acceptable.

### 2.2.4 The data track

The second track in the business dimension lifecycle diagram is the data track and is divided into three phases. Dimensional modeling is the first phase and since the business requirements are stated when reaching the dimensional modeling, the data needed by the users is already known for the developers. The challenge in the dimensional modeling phase is to design a data model that supports these requirements. Inmon et al. (1997) give a rule of thumb with respect to the dimensional modeling and state that the dimensional model of data warehouse should be based in the corporate data model, i.e. the source systems. This will guarantee that the dimensional model will catch the general needs of information in the organization. The best way to do this is to construct a matrix, which contains the key business processes. This matrix will insure that the data warehouse developed is extensible through the whole organization.

When the matrix is developed, it is possible to create a more detailed analysis of operational sources that are going to be used by the data warehouse. This detailed analysis of data sources together with the business requirements definition will give the information needed to construct the dimensional model for the data warehouse. The final product of the dimensional modeling is a logical database model, which will work as a starting point when the physical database is created. The concept of dimensional models will be further discussed in chapter 2.3.2.

The second phase at the data track in the business dimension lifecycle diagram is the physical design that focuses on defining the physical structures that shall support the dimensional model, which is the logical database design. The physical design of the data warehouse also includes setting up the database environment. Another thing that is settled in this phase is the indexing and partitioning of the physical data warehouse.

The third and final phase in the data track concerns the loading of data from the source systems into the data warehouse. This phase is called data staging and development and
phase probably is the most underestimated task in a data warehouse project. This reasoning is supported by Agosta (2000), which claims that the operational costs associated with a data warehouse are five times the cost for the hardware and the software. There are three major parts involved in data staging and development. These parts are also discussed in a more technical context by Chaudhuri & Dayal (1996) in chapter 2.3.1. The first part is the process of extraction, which reveals the quality of the data in the source systems. The data staging and development phase is appropriate for addressing any problems with data quality and possible corrections. The data quality in the source systems is critical in order to make the data warehouse useful. The other two parts are transformation, i.e. cleaning, integration and aggregation, and loading of transformed data into the data warehouse.

2.2.5 The application track
The third track is the application track, which aims to develop applications for the actual users of the data warehouse. The first phase in the application track is the specification of user applications. Inmon et al. (1997) stress the importance of specifying appropriate access tools and claims that the best way of measuring success for a data warehouse is the degree to which end – users find information in the data warehouse, and pass this information on to the organization. The users of the data warehouse can be divided into two categories. One category includes the more advanced users, which are relatively few and performs ad – hoc querying, which in turn requires a bit complicated access to the data warehouse. The other category includes the more ordinary business users, which does not perform the same amount of ad – hoc query processing. A recommendation is that a set of standard end – user applications is developed for the business users, since it is uncommon that they will need any other special features. The application specification will ensure that the developers and the business users have the same understanding of the intended end – user application. Inmon et al. (1997) make the conclusion that without good end – user applications, it does not matter how brilliant the architecture or the modeled dimensions are; the data warehouse will still not be successful.

The end – user application development is the second and last phase in the application track. Since the previous phase in the application track was the specification of end – user applications, this phase concentrates on the actual development of applications based on
2 The data warehouse

given specifications. A good way of developing a standard set of end-user applications is to use advanced data access tools, which will cause a good productivity. The data access tools are used to develop the specified outputs that shall be used by the business users in the organization. Another advantage with using advanced data access tools, except productivity gains, is that the end-users themselves may easily modify and develop existing report templates, in order to get better information that is not specified in the end-user application specification.

2.2.6 Project management
In parallel with the business requirements definition, data track, application track and technology track, project management is running. The most important feature of project management is to make sure that the activities in the different tracks are running in synchronization with respect to the business dimension lifecycle diagram. The project management activities are in use during the whole life cycle. The activities primary mission is to:

“[…] focus on monitoring the project status, issue tracking, and change control to preserve scope boundaries.”

(Kimball et al. 1998, p. 37)

Another important feature of the project management is to establish and maintain a good communication between the business and information systems organization. Without this communication it is hard to reach the data warehouse goals.

2.2.7 Deployment
After the three different tracks have been fulfilled, the deployment phase takes place. It is very important that the tracks for data, technology and applications can work together, since they together build the data warehouse. Before any users get access to the data warehouse, detailed planning must be made. This planning contains education of business users, user support and strategies for feedback. Finally the deployment phase should be stopped, if it reveals that any of the different tracks are not ready for release.
2 The data warehouse

2.2.8 Maintenance and growth
When the data warehouse finally has been deployed, is up and running, and used by business users in the organization, the data warehouse need to be maintained in different ways. One part of the maintenance is to perform ongoing support and education for the business users. It is very unlikely that all business users in an organization adapt all the features provided by the data warehouse at once. Continuous education and information of new features and updated functions is useful in order to give the business users good possibilities to use the data warehouse in an efficient manner. Another important aspect of the maintenance of a data warehouse is the actual organization. The organization uses business processes, which stated earlier, constitutes the foundation for the data warehouse. If processes in the organization are changed, deleted or added, it is important that the data warehouse is updated with this new information, otherwise the data warehouse will not be efficient. It is important to have a long term plan for the continuous development of the data warehouse. The plan will ensure that new needs from business users are adapted and that there is a good communication between different parts of the organization. This statement is highlighted by Haley (1998) who states that requirements change and is developed over time. The reason for this is that more and more users understand the possibilities with the existing data warehouse. Before the data warehouse actually exists, i.e. during the beginning of a data warehouse project, it might be hard for every user to specify his or her needs. This is important since the business dimension lifecycle diagram does not have an ending, just a back – loop to the first phase, which is project planning, which means that the cycle is a never ending process of improving an organizations data warehouse.

2.3 Data warehouse architecture
Since this project is aimed to investigate the relation between business processes and data warehousing, the focus will not be at the technology behind data warehouse. However, it is still necessary to briefly explain the underlying technology of data warehousing. The reason for this is to clarify the position of the data sources, which offer data to the warehouse, and the role those tools offer to extract knowledge from the data warehouse. The data sources are considered important, since these systems support the business processes, which the data is collected from. We intend to go through and explain all the steps in a typical data
2 The data warehouse

warehouse architecture provided by Chaudhuri & Dayal (1996) in order to introduce the concept of data warehousing in a broader manner to the reader.

Chaudhuri & Dayal (1996) present a figure of an ordinary data warehouse architecture. The framework represented is a familiar representation and is used (a bit modified) by numerous authors, e.g. Connolly et al. (1999), Breitner (1997) and Kimball et al. (1998).

![Data warehouse architecture](image)

Figure 3: Data warehouse architecture. (Chaudhuri & Dayal, 1996, p. 66)

2.3.1 Operations and tools in the data warehouse

The data warehouse architecture is equipped with tools for collecting and extracting data from the various sources and integrating it into the data warehouse. The loading of data into the data warehouse consists of four different operations, according to Chaudhuri & Dayal (1996):

- Extraction is a process that concerns what data that shall be stored in the data warehouse. Only data that can be used for decision making and other typical data warehouse operations is of interest. Kimball et al. (1998), also support this view of extraction process.
- The transformation of extracted data is the next step in order to create a data warehouse. It is necessary to transform the data since its common that data contain different errors, such as different field lengths and inconsistencies. There are different operations related to the transformation step. Data cleaning is the
correction of misspellings and broken integrity constraints. **Aggregation** of data means that data is stored at different levels of detail and this improve the performance of the data warehouse when the user runs common queries. **Integration** is the final operation in the transformation step and this means that data on different formats is viewed in one ordinary format that users apply when accessing the data warehouse.

- After the data has been extracted and transformed, the clean, consistent and aggregated data is loaded into the data warehouse. Taking the data warehouse offline makes the load of data easier, and this is usually done at night, when the data warehouse is not in use.

- The last operation mentioned by Chaudhuri & Dayal (1996) is refreshing the warehouse. Usually, the warehouse is refreshed periodically, i.e. every change in the source systems is not updated at once, but is saved and loaded with other changes that have occurred during a particular time interval.

In order to handle data in the warehouse and to transform, load and refresh the data from its sources, the architecture provided by Chaudhuri & Dayal (1996), contain a repository for meta data storage. Connolly et al. (1999), state that the repository stores meta data definitions, used by all the processes in the warehouse. According to Breitner (1997), meta data provides information about the origin, actuality and quality of the data. Kimball et al. (1998) describe meta data as all the information in the data warehouse that is not data.

### 2.3.2 OLAP servers

Once the data from the different source systems have been loaded into the data warehouse, it must be stored. Chaudhuri & Dayal (1996) claim that the most frequent way to store data is to use the multidimensional model. Kimball (1996) describes the multidimensional model such as a cube. Chaudhuri & Dayal (1996) describe the multidimensional model such as a source for numeric measures. An example of measures may be sales and inventory. Every numeric measure is dependent of a set of dimensions which, according to Chaudhuri & Dayal (1996), provide the context for the measure. The dimensions in
2 The data warehouse

combination are then determining the value of the measure. We want to clarify this by illustrate how Chaudhuri & Dayal (1996) describe multidimensional data (Figure 4):

Figure 4: Multidimensional data (Chaudhuri & Dayal, 1996, p 68).

This multidimensional model presents how the data warehouse data is stored on the OLAP servers.

2.3.3 Data marts

Building a data warehouse is a long and complex process that requires a lot of resources and, according to Chaudhuri & Dayal (1996), some companies are settling for data marts instead. Watson et al. (2001) support this statement, by saying that organizations usually begins to create one data mart and then expand to a data warehouse later on. Data marts are subsets of a data warehouse, which are related to a certain section of a company. A data mart contains those subjects that are interesting for a particular section. Chaudhuri & Dayal (1996) say that data marts are a good solution, since they are built faster. The reason for this is that it is not necessary to consider the whole organization, only the parts that are of interest for the particular section has to be considered. Another feature mentioned by Connolly (1999) is that data marts contain less information than a data warehouse. This means that data marts are easier to understand and navigate than the entire data warehouse. The drawbacks with data marts are that it may be complex to integrate different data marts, since an incomplete model of the company is used (Singh, 1998). Another drawback mentioned by Watson et al. (2001), is that a data mart only supports a limited number of users.
2.3.4 Front end tools

According to Chaudhuri & Dayal (1996), front end tools make the data in the warehouse available to the users of the warehouse. Front end tools consist of data access and retrieval tools and can be categorized in five categories according to Connolly et al. (1998):

- Reporting and querying tools are tools that include production reporting tools that generate and report progress in operational systems. Querying tools are tools that handle simple SQL queries, which are run in the data warehouse.

- Application development tools are tools that can be used to create parts of a query tool or refine some standard tool by the user if it is necessary.

- Executive information systems tools are used to support management decision making at all levels in a company. The user may modify the application to its own personal needs, in order to make better and faster decisions.

- Online analytical processing (OLAP) tools allow a user to analyze data using complex, multi-dimensional views. Inmon et al. (1997) state that OLAP tools make it easier to perform multidimensional analyzes. This kind of tools takes for granted that the data is organized in a multidimensional model. Chaudhuri & Dayal (1996) describe dimensions as a context for a measure. For example, the dimensions city, product name and date may be associated with the measure sales. Chaudhuri & Dayal (1996) mention another interesting feature about the conceptual model for OLAP. This is aggregation, which means that a certain dimension may be viewed in different levels of detail.

- Data mining tools are characterized by Chaudhuri & Dayal (1996) as the search techniques that are used to detect information that is hidden in the data in a data warehouse. According to Söderström (1997), data mining tools are tools that may be learning systems that can perform analyzes and model building. Connolly et al. (1999) take this reasoning a bit further and claim that data mining is a process that shall result in discovering interesting, new correlations, patterns and trends. Analyzing large amounts of data stored in a data warehouse or a data mart does this. Connolly et al. (1999) and Singh (1999) state that data mining tools are suitable for
2 The data warehouse

decision support because of their ability to visualize the result of a query, in a clear and broad way.

The introduction of front end tools has, according to Chaudhuri & Dayal (1996), made it possible for non expert users to access the data warehouse. Often, query environments are offered, that help the user building complex ad hoc queries by “pointing and clicking”, which makes queries easy to handle (Chaudhuri & Dayal, 1996).
3 Business Processes

3 Processes

A process is, according to Rentzhog (1998), a sequence of activities that creates value for a customer. Davenport (2000) uses the same basic ideas when reasoning about processes and states that a process consists of a specific order of activities with a beginning and an end. These activities must have clearly defined input and output. Eriksson & Penker (2000) use models to describe processes. According to Eriksson & Penker (2000), a process always has a goal that motivates the process. Furthermore, every process has an input and an output. The process is then complemented by different resources and activities, which support the process of moving an object from input to output. The main difference between these statements is that Rentzhog (1998) mentions the customer. By adding the customer into the concept of process, we are narrowing different definitions of the concept business process.

From this reasoning the conclusion may be drawn that a general process concerns activities that shall produce something. By adding a customer we are moving from the general process to a business process. In order to make the concept of a process more apparent we present a figure that covers the basic ideas about a business process (Figure 5):

![Diagram of a business process]

Figure 5: Description of a business process and fundamental concepts, derived from Eriksson & Penker (2000, p. 69).
3 Business Processes

3.1 Business processes
According to Kimball et al. (1998), a business process is a set of activities, intended to create a value for a customer. This view of business processes is commonly used, and is supported by Rummler & Brache (1995), Willoch (1995), Davenport (1993) and Harrington (1991). Business process are often overlapping with other business processes, and as a result of this, Kimball et al. (1998) choose to view a business process such as a helpful grouping of information resources with a consistent theme. In order to illustrate a business process, Eriksson & Penker (2000) use activities inside the business and show how the business process narrates to and cooperate with resources in the business. This is done to accomplish a certain aim for the business process. Franke (1999) gives a bit more general description of business processes, and claims that a business process simply characterizes the way things are done in an organization.

As we already stated, the main difference between a process and a business process is the attention to the customer. From this we can draw the conclusion that a business process is a special kind of process, which is used to create value to a customer. From the above statements of business processes we have created an operational definition that covers the aspects of this thesis:

“A business process is a sequence of coordinated activities, which creates value to the customer. By using resources, a process is transforming input to output, aimed for an external or internal customer”

The definition given focuses on presenting that a business process is a process that creates value. An important aspect of the definition is that the customer may be internal or external. This is important since everything produced in an organization is not aimed directly to an external customer who pays the organization money. The customer may in many cases be an actor in the organization who needs a service of some kind.

According to Willoch (1995), there are some characteristics of business processes. We have already mentioned that we view a business process as a certain type of process directed
3 Business Processes

towards a customer. The characteristics for a business process are, according to Willoch (1995):

- Business processes are not functional specific, i.e. they concern more than one function within an organization.
- There is no one in charge for the business process. Since the business process concern more than one function in the organization, it is hard to find any person or department responsible for the business process. Instead, every function is responsible for its part of the process.
- Business processes do not have clear names within the organization. The reason for this is that an organization is often described by its departments and that a process often concerns more than one department. Due to this, it is difficult to assign a process to a specific department.
- Business processes are invisible. This stems from the fact that an organizational schema often just describes which person that is responsible for what area in the organization. There is no description how the processes flow through the organization.

These characteristics show that business processes are cumbersome to handle, since they are hard to make clear and explicit, when analyzing an organization.

3.2 Different types of business processes
Rentzhog (1998) states that it is possible to divide or categorize business processes, with respect to their function in an organization. The following categorization is given by Rentzhog (1998):

- Strategic processes that are used to create plans for the organization and build the foundation of the organization.
- Operative processes that generate a direct value for the customer. It is the operative processes that create outputs to the customer and revenues for the organization.
3 Business Processes

- Supporting processes that do not create direct incomes to the organization, but supports strategic and operative processes. Supporting processes are administrative and technical processes and their support of other processes allows for a well working organization.

This categorization is also supported by Rosander (1997). Rummler & Brache (1995) give a similar categorization, but with a slightly different focus. The main difference between the two categorizations is that Rummler & Brache (1995), in a more explicit way, mention if the customer is internal or external. Another difference between the categorizations is that they are named differently. The categorization given by Rummler & Brache (1995) is the following:

- Core processes, which results in a product or a service that is aimed for an external customer. Examples of this kind of processes are development of products and services.
- Supporting processes are processes with an output that is invisible for the external customer. Examples of this kind of processes are processes that are needed in order to fulfill the core processes, such as education of staff.
- Management processes aims to support the business processes in an organization. This may be strategic and tactical planning.

The categorizations of business processes made by Rentzhog (1998) and Rummler & Brache (1995) make it clear that there is a rather common understanding of the consisting elements of business processes. We have chosen to use one uniform way of referring to different types of business processes, which will be used from now on in the thesis. We chose to use the categorization made by Rummler & Brache (1995) since their categorization explicitly states in what category of business process that the customer is internal or external. We also believe that the categorization given by Rummler & Brache (1995) is clearer in general and thereby easier to understand.
3 Business Processes

We have already stated that a business process consists of one or more activities. It is thus possible to decompose a business process into one or more processes. Rentzhog (1998) states that every business process consists of a number of participating processes, which in turn consist of activities. An activity is, according to Bergstrand & Wallin (1995) the driving force in the process. Bergstrand & Wallin (1995) distinguish between different types of activities. Hard activities concern physical things such as transportation. Soft activities are less explicit and concern areas such as research and planning.

In this thesis we will mainly focus on different types of business processes and we will not be concerned with any further decomposition of business processes. The argument for this is that at a lower level (instance level), business processes are named very differently and it would be very hard to compare and analyze the collected material.
4 The interviews

As was stated in the Introduction (chapter 1.3), we will conduct interviews to gather the information needed in order to fulfill the aim and objectives of this thesis. The interviews will be performed by telephone, since the respondents do not have the time available to sit down in a face-to-face interview. This chapter aims to give a brief description of the intended questions that will be used during the interviews. We will also discuss how we intend to analyze the information gathered during the interviews.

4.1 The interview questions

We intend to create a number of basic questions that will cover the fundamental aspects of the project. As a complement to the basic questions, we will create a number of questions that describe the respondents’ experience of data warehousing and also how the respondents view a data warehouse. We intend to keep the interviews at a conversation level, since we consider it to be important that the respondent has some space to follow up the basic questions. Patel & Davidson (1994) call this kind of interviews high standardized, which allows the interviewer to be more flexible and discuss certain details at a more general level if necessary. A high degree of standardization is appropriate when the purpose of the interview is to compare and generalize information from different respondents.

The questions that are going to be used will have a low degree of structure. The degree of structure is, according to Patel & Davidson (1994), how much freedom the respondent has when answering questions. A low degree of structure will increase the possibilities of interpretation that the respondent has when answering questions, which will give the respondents a lot of space when answering the questions.

4.2 Analysis of collected material

The categorization of business process types that are more commonly affected is thought to be made quite straight forward. We have in previous sections (3.2) stated different types of business processes. Our intention is to sort those business processes that have not been categorized by the respondents, into the three types of business processes mentioned in chapter 3.2.
4 The interviews

The factors that trigger the need for changing business processes need to be categorized in a more extensive way. We have in studied literature not found many data warehouse specific factors that cause a need to change business processes. The only thing mentioned is that need for data in the data warehouse cause a need for changed business processes. We want to investigate these factors further in order to reach a more extensive picture of which factors that cause these needs for changing processes.

Finally, we intend to find out if there is any correspondence between certain factors and certain business processes affected. This will be done after the other two objectives have been fulfilled since the outcomes of these are necessary in order to establish any correlation between them. Our goal is to create a form of framework where it is possible to make a clear correlation between factors and business processes.
5 Interview results

This chapter gives a presentation of the material gathered during the interviews. The material will be presented in the following manner. We intend to go through the answers from those questions that describe the respondents, their experience and their roles in data warehouse projects. Furthermore, we will summarize the information caught in those questions that is directly related to the objectives of this thesis, i.e. what types of business processes that are usually affected when introducing a data warehouse project in an organization, what factors there are that triggers this need for changing the business processes during the pre-deployment stage of a data warehouse project. We will also present information that concerns the correspondence between categories of business processes and factors that triggers a need for change in these business processes.

5.1 The respondents

The respondents are all employees at well-known system development companies. With one exception, the respondents are working as data warehouse consultants and implements data warehouses into other organizations. Only one respondent (respondent 6) is working with his company’s own data warehouse. This means that all respondents, except one, have been participating in data warehouse projects as external consultants. The tasks these external consultants have had are mostly analysis of organizational requirements, design, and implementation of data warehouses.

The respondents’ experience of data warehouse development is to consider as thorough, as the average years of working with developing data warehouses are 6 years, and the fewest years working in the area are two. Furthermore, all respondents have experience from other types of ISE, and several of the respondents have worked in parallel with other types of information systems. We have chosen to illustrate each respondent’s experience (in years) in Table 2.
The number of respondents that have been interviewed is 8. In order to reach suitable respondents, a data warehouse consultant was used as a contact person towards the respondents. The contact person introduced the thesis aim and objectives to colleagues and those who were interested in participating in interviews were identified. Contacts were made with interested actors, and the majority of these were interviewed. Not all actors were interviewed because they did not have the time available during those weeks when the interviews were going to be conducted. The interested actors that did not get interviewed were 27% of the total amount of interested actors.

All respondents, with one exception (respondent 2), have been involved in more than one data warehouse project. The average of data warehouse projects that the respondents have been participating in is 10. This is considered as contributory, since the respondents by that have a substantial experience of data warehousing projects and may thereby answer the interview questions from an expert’s perspective. We also argue that 10 data warehouse projects are extensive, since the number of data warehouse developers in Sweden is limited and each project often consumes a lot of time. We have chosen to illustrate the respondents experience (in projects) in order to give a more clear view (Table 3).
5 Interview results

Table 3: The number of DW projects that the respondents have been involved in.

<table>
<thead>
<tr>
<th>Number of projects the respondents have been involved in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondent</td>
</tr>
</tbody>
</table>

The majority (63%) of the respondents also have a technical background, which means that they all have been working with the technical parts of implementations of data warehouses. However, the respondents current roles are more focused on the coupling between the organization and the data warehouse, which means that the respondents has a good understanding for the organizational aspects of data warehouse development as well as the techniques behind data warehousing. The percentage distribution of the respondents’ roles is given in Table 4 (Each respondent may have held more than one role).
Table 4: The respondents’ roles in DW projects

5.1.1 The respondents view of a data warehouse
The respondents generally describe a data warehouse in two ways and some respondents use both descriptions. One type of description concerns the technical aspects of a data warehouse. This description fits well with the architecture provided by Chaudhuri & Dayal (1997) and is mentioned by all respondents. The other type of description mentioned by the respondents is the actual use of the data warehouse and the benefits with adopting a data warehouse, which is mentioned by respondent 1, 3, 4, 6, 7, and 8. One benefit described is for instance correct information that is gathered from the organizations source systems, which improves the business decisions in the organization. All respondents mention that a data warehouse should generate some form of increased value for the organization.

From this we conclude that the definition given in this thesis fits rather well with the descriptions mentioned by the respondents, i.e. a set of databases that makes it possible to manage data that is used to be a support for tactical and strategic decision makers, and thereby generate faster and better decisions.

5.1.2 The respondents view of a business process
The respondent’s general view of a business process (all respondents) is that the business process should generate some kind of value for the organization. Some of the respondents (3 & 4) view a business process as a flow that increases the value for an input. Some of the
5 Interview results

respondents also mention the involvement of a customer (4, 5, 6 and 8), but none of the respondents distinguish between an external or internal customer. The respondents often exemplify a business process with manufacturing processes, sales processes or purchasing processes.

The descriptions of a business process given by the respondents differ a bit with the definition given in this thesis. The main difference is that the respondents do not include the intervention of an external or internal customer in the same way as the definition given in the thesis. We still think that the descriptions given by the respondents fit rather well with the definition, which means that the respondents have discussed business processes within the context of this thesis.

5.2 Types of business processes affected

The respondents have made a clear distinction between post and pre-deployment in their answers. The ideas from the respondents are unified: It is highly common that a need to change business processes is detected during the pre-deployment stage of a data warehouse project. However, the need detected does not always result in a direct change of business processes at the pre-deployment stage. The reasons for this will be further analyzed in chapter 6.

As a reminder we once again state the three categories of business processes that are discussed in chapter three:

- Core processes, which results in a product or a service that is aimed for an external customer. Core process may be exemplified by manufacturing processes and some service processes.
- Supporting processes which are processes with an output that is invisible for the external customer. Supporting processes are needed in order to fulfill the core processes and may be exemplified by service processes such as education, accounting, and relations to customers and suppliers, i.e. logistics.
5 Interview results

- Management processes are processes that aim to support other business processes in an organization, i.e. core and supporting processes. Processes that concern strategic and tactical planning may exemplify management processes.

From the interviews, we put together the core essence from each interview. Some respondents only discussed business processes at a type level, which made these interviews easy to summarize. Other respondents discussed business processes at an instance level. When this was the case, the business processes discussed were classified and placed into one of the three categories available. The frequency of discussed business processes is given in Table 5.

All respondents state that supporting processes and management processes are the types of business processes most often affected by a need for change. As stated before, supporting processes are those business processes that have an output that is invisible for the external customer, i.e. processes that are needed in order to fulfill the core processes. Examples of supporting processes, mentioned by the respondents, are different kinds of logistic processes and account processes. The majority (63%) of the respondents also mention management processes. The management processes are, as stated in section 3.2.1, processes that support other processes in the organization. Examples of management processes are processes that concern marketing and care for customers, i.e. strategic and tactical planning. Core processes are also mentioned by some of the respondents (4, 5, and 7) but are considered hard to actually change at the pre-deployment stage of a data warehouse project compared to support and management processes. We will discuss the reasons for this in chapter 6.

Table 5: Frequency of given types
5.3 Triggering factors that may cause a need for change

A general agreement among the respondents is that it is crucial to change the business processes, in order to be able to achieve the full potential of the data warehouse. An example of this is the production of correct (without errors) data in the source systems, which will be used by the data warehouse.

Another triggering factor mentioned (by respondent 1, 3, 5, and 8) is that the data warehouse development project itself triggers a need for change. What is meant is that the data warehouse makes the organization aware of new aspects that may approve the business processes in the organization.

One triggering factor mentioned by the respondents (1, 4, 5, 6 and 7) is that new knowledge of the organization, which appears during analysis and requirements specification, results in a need to change some business processes that does not behave the way people think.

One further triggering factor mentioned (by respondent 3 and 6) is that the introduction of a data warehouse makes the organization aware of primary goals, for instance increased degree of service for customers and suppliers, which in turn creates a need to change or correct some concerned business processes.
5 Interview results

Respondent 2 is missing in the previously mentioned triggering factors. The reason for this is that the respondent could not give an answer related to the pre-deployment stage of a data warehouse project. Triggering factors that concern the post-deployment stage will be discussed at a general level in chapter 7.

Table 6: Factors that might trigger a need for changing business processes

<table>
<thead>
<tr>
<th>Factors that might trigger a need for change</th>
<th>Amount of respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Achieve maximum ROI</td>
<td>80%</td>
</tr>
<tr>
<td>DW project itself</td>
<td>70%</td>
</tr>
<tr>
<td>New organizational knowledge</td>
<td>60%</td>
</tr>
<tr>
<td>Focus on primary goals</td>
<td>50%</td>
</tr>
</tbody>
</table>

5.4 Correspondence between triggering factors and business processes

During the interviews no general correspondences between certain business processes that need change and triggering factors has been found. Two different categories of answers have been given from the respondents. One group of respondents (4, 5 and 6) states that there does not exist any general correspondence. Triggering factors and affected business processes are project specific. The other group of respondents (1, 2, 3, 7 and 8) cannot give an answer.

Although, we have identified some types of correspondence at a general level. The respondents’ answers to other questions during the interviews make it possible to implicitly derive some general correspondences between triggering factors and concerned business processes. These correspondences will be discussed and analyzed in chapter 6.
6 Analysis

This chapter will analyze the material presented in the previous chapter. The analysis will also be complemented and contra stated with ideas from related literature.

6.1 Types of business processes affected

The unanimous answer from the material collected was that it is common that a need to change business processes arises in the pre-deployment stage of a data warehouse project. Supporting processes are mentioned by all of the respondents, a majority has mentioned supporting processes and some respondents have mentioned core processes. Although, there is a difference in handling different types of business processes when it comes to the actual change of these. Core processes are considered harder to change at the pre-deployment stage compared to supporting and management processes.

6.1.1 Core processes

The main argument for not changing core processes at the pre-deployment stage is that they often has a central position in the organization, especially when the organization is producing something tangible, such as different components etc. As a result of this, core processes are often considered to be too important for the organization, and the alignment with the data warehouse comes second for the organization. Another reason why core processes are considered hard to change is that core processes are often well documented and have been optimized over time, which often results in that an organization does not believe that there really exist a need to change this category of business processes. Dubé & Robey (1999) discuss this problem and states that a major trigger for this dilemma is uncertainty. If actors in the organization only know that some processes are going to be changed, without knowing how the changes will be executed, people in the organization will not accept these changes. By reducing uncertainty about the future, the actors will experience a sense of control on their working environment, and it will be easier to introduce suggestions for changes also in the core processes.

Another issue is the people working in the core processes. Many of the respondents mention that it is quite easy for an external consultant to detect those processes that need change. It is another thing to introduce this to the actors in the organization. A widespread
opinion among the respondents is that it is hard to convince actors in the organization that the processes they are involved in need to be changed. The actors in the organization must discover this by themselves and this discovery cannot be done until the data warehouse is up and running, i.e. in the post-deployment stage of a data warehouse project. This reasoning is supported by Harker et al. (1993), who state that the introduction of any type of information system into an organization unavoidably results in the finding of new ways of doing things and new knowledge that reveals new possibilities. This type of effects are, according to Harker et al. (1993), often found once the users of the system have experience from using it, which in this case is the same as the post-deployment stage. According to Bubenko (1993), no system can be built in isolation. The system must always be related to both user requirements and existing systems, in this case the source systems. Bubenko (1993) also underlines the importance of doing necessary changes in early phases, i.e. the pre-deployment stage, since changes in an already operating system consumes a lot of resources. Respondent 3 gives one solution to the problem, which is ideal but probably hard to actually carry out, and states:

“It is important to get rid of suspicions in a company and kill the myth that all processes are holy.”

(Author’s translation)

The fulfillment of this statement is, according to collected information, not an easy task. But if the problems remain unsolved it might result in that the core processes concerned, which are not optimized, might produce bad or even worse, incorrect data that is collected and integrated into the data warehouse. The reason for this dilemma might be that the actors in the core processes sense an ownership to their business processes, and they do not want to change their way of doing things. One explanation to this behavior, mentioned by Nilsson (1997), is that the change of core processes, as well as other types of business processes, involves change of concepts, which is the core of the organizational structure. By changing the business language, i.e. the concepts, the mental models of how thing are done must be changed. This is usually controversial and such a process has to evolve over time.
Another explanation to this problem can often be deduced from the early stages in the business dimension lifecycle described by Kimball et al (1998), i.e. business requirements definition and dimensional modeling, which have been discussed in chapter 2.2. The problem arises when actors in the organization do not feel a commitment to the development project, and by this, do not participate enough in the process of development. A possible solution to this problem is, according to Harker et al. (1993), to encourage the users to actively engage themselves in the project. This will, according to Ang & Teo (1993) and Harker et al. (1993), result in the fact that the users experience a better control over their working situation. Cherry & Macredie (1999) use a similar solution when adopting participatory design, which promote a sense of ownership and offer users a vested interest in the success of the system. This will probably result in a higher acceptance of the system when it is implemented in the organization. Strehlo (1996) has the same belief and comes to the conclusion:

“*If you involve users in each design phase of your warehouse and then seek out other potential users when you roll out each phase, they’ll come back to use the system.*”

(Strehlo, 1996, p. 7)

One other reason why actors in an organization are unwilling to change core processes before the data warehouse is up and running is, according to the information gathered during interviews, that it is common that a data warehouse project focuses too much on the technical aspects. Cherry & Macredie (1999) state that users are experts in their own working environment and have unique insight into problems and possibilities in this environment. Gray & Watson (1998, p. 11) describe this issue and states:

“The business issues – sponsorship, identifying business needs, politics, training and supporting users – are critical. Although data warehousing professionals verbally recognize the importance of these issues, all too often they revert to focusing on the technical issues. They need to better accept the need to wear both technical and business hats.”
By using expert knowledge actors in the development project, Cherry & Macredie (1999) state that unnecessary or redundant processes can be changed or removed easier than the case would be if the users were not participating. If there is too much focus on technical aspects and too less attention is directed towards organizational aspects, which the actors in the organization understand and are experts on, we will most likely experience the same problem as previous, i.e. a sense of ownership among the users that has a negative influence on the data warehouse project. A too strong emphasis on technical aspects may also, according to Strehlo (1996), cause interest among IS staff in the organization, but no other will bother to care. If implementers are unaware of the importance of stressing both technical and organizational aspects, their chance of success will be reduced.

As already stated, core processes as well as supporting processes and management processes, are highly likely to experience a need for change. But to actually change core processes is, as mentioned previously, much harder than changing the other types of business processes discussed. We will therefore further discuss supporting processes and management processes, since these are considered easier to actually change at early phases of the development without encountering the same amount of problems as for the core processes.

6.1.2 Supporting processes
Supporting processes is one type of business process that is mentioned by all of the respondents as often being exposed to needs for change. Data produced in this type of business processes are often used as an input to the data warehouse, and is therefore critical, with respect to high quality data, in order to get a data warehouse with good performance and trustable output.

The most frequently mentioned examples of supporting processes pointed out by the respondents are processes that concern relations with customers and suppliers, such as logistic and accountant processes. These processes do not concern the actual production of things in the organization. They are needed in order to get input to the organization, and
selling the output to customers and thereby making money, which is the primary goal for the organization.

Since supporting processes are intended to offer support to the core processes in the organization, the respondents view these as easier to change. The arguments for this is that supporting processes are not in the center of the organization, which makes them easier to change without affecting the production of goods or services in a straight forward manner. Another argument mentioned by the respondents is that supporting processes are often less optimized and documented, which may result in a need for better performance and documentation. Still, we might experience the same problems with user acceptance within this type of business processes.

6.1.3 Management processes

The management processes are the second type of business processes mentioned by the respondents (63%) as likely to experience a need for change. Some of the respondents state that it is the management processes that form the foundation for the organization. Without management processes, no other business processes are needed since the organization would not exist. What the respondents mean is that if the organization does not have any business processes that are related to marketing, strategies etc., the organization would not have any customers and thereby the organization would not last long at the market. Respondent 3 states:

“[…] a core process can not be improved by itself; you have to start the improvement at the management level, since everything in the organization is directed to sell a certain product to a customer. If you do not have any customers, it does not matter that your core processes are excellent.”

(Author’s translation)

The definition given of a data warehouse in this thesis implies that we struggle for increased organizational value, i.e. better knowledge about the organization itself and its environment. This new knowledge is intended to result in increased ability to sustain competitive edge by targeting new customer segments and finding new, more efficient
ways of selling products or services. With this view of data warehousing, there is a strong relation towards management processes, since this category of business processes concerns the support of other business processes in an organization, such as strategic and tactical planning. Strategic and tactical planning is, according to Strand (2000), performed at strategic and tactical levels in the organization. These levels are using information systems such as executive support systems and management information systems, which do not concern the actual doing of things, and differ from OLTPs, which is used at the operational level in the organization. The introduction of a data warehouse will result in a new way of viewing the business processes that are related to information systems used at higher levels in the organization and will probably result in improvements of how things are done. The actual work at the strategic and tactical levels may also be improved by the introduction of a data warehouse, since a lot of work related to strategies and tactics may be performed by using the data warehouse instead of using the existing computerized systems and techniques. From this the conclusion can be made that it is obvious that the introduction of a data warehouse project give rise to a need to change and also add some new features about the way work is done at the management level when a data warehouse project is introduced.

6.2 Factors that may cause a need to change business processes

There are many factors that may trigger a need to change business processes in an organization and a majority of these factors are not specific to data warehouse development, but are rather common to ISE in general. The introduction of an information system in an organization works, according to the information collected, as a catalyst, which always results in a need to change how some things are done. The development of information systems forces actors in the organization to view their way of working from a new angle with a possible consequence of new organizational knowledge, which often results in a need to change the business processes concerned. Nilsson (2000) describes the relationship between the reality and the information system. Both reality and information system affect each other and the introduction of an information system will automatically result in changes in the organization but the reverse must also be recognized. Cherry & Macredie (1999) discuss the same topic and state that if changes occur in the organization, it is obvious that these changes are also reflected by the information system, since the
information system is a simplified representation of the real world, within the context of the organization.

The factors that suggest a need for changing business processes can roughly be divided into three levels of abstraction. At a general level, triggering factors that affect all types of business processes are similar to other types of ISE, but because of high development costs and complex environment, data warehouse projects are stronger related to return on investments, data quality, control of data etc. At a type level, there have been identified some correspondences between types of business processes affected and their triggering factors. We have excluded the core processes, because of the difficulties with changing these at the pre-deployment stage, although, the correspondences presented may be applicable to core processes as well. At an instance level, no correspondence has been found since business processes are organization specific, such as process names and participating activities, which also make business processes unique for a certain development project. We have chosen to illustrate this in figure 6.

<table>
<thead>
<tr>
<th>General/ISE level</th>
<th>Suitable for all types of business processes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type level</td>
<td>Specific to certain types of business processes.</td>
</tr>
<tr>
<td>Instance level</td>
<td>Specific for each organization and project.</td>
</tr>
</tbody>
</table>

Figure 6: Levels of triggering factors

6.2.1 General level
The overall general factor that causes a need to change an organization’s business processes is, according to the gathered information to achieve maximum return on investments from the data warehouse. Such change may for instance affect the production of the required data in the source systems.
Another general triggering factor, that is commonly mentioned, is that the data warehouse project makes the organization aware of things that would improve the business processes. These things are not necessarily related to the data warehouse, but may be highlighted anyway within the data warehouse project. This awareness of things that has not yet been observed may be a result from the analysis of the organization or from the requirements specification.

New organizational knowledge is another general triggering factor. Since a DW project requires analysis of the organization in order to state the DW requirements, actors may be forced to reconsider how things are executed at present and what changes that are needed in order to get to a desired future state.

Another general factor mentioned in the material is that the organization may need to reconsider its primary goals. What is meant is that the customer, i.e. the organization, is being forced to change priorities between its primary goals with respect to the data warehouse. Respondent 3 states:

“The triggering factors are mainly the definition of values and the efficiency that is demanded in order to get from present to desired future state.”

(Author’s translation)

When an organization prioritizes its primary goals, it automatically pays more attention to business processes that are related to the primary goals. This will probably result in a need to change some business processes, since the way things are done in the organization must be changed, in order to get from present state to future desired state.

### 6.2.2 Type level

Objective 3 of this thesis concerns the investigation of correspondences between the affected business processes and factors that trigger a need for change. At a type level, some correspondences between affected business processes and triggering factors have been identified. Core processes have been excluded, in view of the fact that these are considered...
much harder to change than supporting and management processes at the pre-deployment stage. However, we are aware that the correspondences presented are applicable to core processes as well.

Correspondences between supporting processes and their triggering factors, found in collected material are:

- A determination in the organization to improve the way of working with customers and suppliers. As already mentioned, one of the arguments for developing a data warehouse is to gain better knowledge about customers and suppliers. Some supporting processes, for instance logistics, are directly affected by an organization's ambition to improve its way of working with customers and suppliers.

- The need to produce the required data in the source systems. Data collected and integrated into the data warehouse often stems from the organization's transaction systems, which mostly concern core processes. However, data stemming from supporting processes is also used by the data warehouse, for instance logistics and the need to produce required data in the source systems will affect some supporting processes as well.

- The need to produce correct (error free) data in the source systems. This triggering factor is, such as the previous mentioned factor, not specific for supporting processes. But the data used by the data warehouse stemming from supporting processes is as important as any to be correct. Because of this, we argue that the need to produce correct data in the source systems will trigger a need to change those supporting processes that are concerned.

Typical management processes in an organization are strategic processes that are related to other business processes in the organization. Examples of management processes are marketing and long term planning with respect to research and development, commercial campaigns etc.
The factors that might cause a need to change organizations management processes during the pre-deployment stage of a data warehouse project are:

- Increased competition, which requires better and faster decisions with respect to marketing, flexible production etc. This factor is also one strong reason to develop a data warehouse in the first place, since the data warehouse is intended to support decisions that result in sustained competitive edge.

- The alignment between processes and data warehouse will increase the opportunity to find new market shares. This is an obvious triggering factor. If a data warehouse is to be well performing it must be properly aligned with related business processes. Nilsson (2000) illustrates the strong relation between organization and information system and emphasizes that changes in one unit will automatically cause changes in the other unit.

- The data warehouse development might result in a new way of viewing the organizations business vision by new priorities etc. The introduction of a data warehouse will make the organization aware if its primary goals. The primary goals might be many and all primary goals cannot be fulfilled, since this would result in a never ending development project. By being forced to prioritize its goals, the organizations have to carefully look into the concerned processes, and perhaps view these in a new way.

6.2.3 Instance level
At an instance level of business processes, no correspondences have been found between triggering factors and affected business processes. Respondent 4 states:

“It is probable that it exist general truths but often, triggering factors are unique for a certain project.”

(Author’s translation)

In view of the fact that business processes are organization specific, which results in that business process names and participating activities in these processes are unique for each
6 Analysis

organization. Since a data warehouse most often is developed within one single organization, the result is that business processes are unique for a certain development project.
7 Result

This chapter concludes the analysis. We intend to go through each objective and finally conclude the aim of this thesis.

7.1 Objective 1

The categorization of business processes was made according to Rummler & Brache (1995) and distinguishes between three different types of business processes; core processes, supporting processes, and management processes.

It has been found that all types of business processes are affected when a data warehouse project is introduced in an organization. This need always appear when any type of computerized information system development project is introduced. Despite a need to change all types of business processes, core processes are not changed to the same extent. There are several reasons for this and we will give a brief summary of the most frequently discussed:

- Core processes often have a central position in the organization and are therefore considered to be too important for the organization, i.e. core processes are in the center of the production and changes in these will cause stops and delays that in turn will cause decreased incomes for the company. This results in that the alignment with the data warehouse gets a low priority.

- Core processes are often well documented and have been optimized over time. This results in that the organization does not believe that there really exists a need to change the core processes. This dilemma is often associated with uncertainty about what changes will result in. By reducing uncertainty it will be easier to introduce suggestions for changes among the core processes.

- Core processes are often well known by actors and it is hard to convince the actors that core processes need to be changed. The reason for this may be that the actors sense an ownership of the processes, which is a negative influence when changing core processes. The actors must experience a need for change by themselves, but this cannot be experienced at the pre-deployment stage of a data warehouse project.
7 Result

- Another reason is that the actors do not participate enough in the development process. By involving actors in the development project, the negative sense of ownership may decrease and it is also highly likely that the finished DWs will show better performance, since expert knowledge about organizational aspects has been used to a greater extent.

- A too strong emphasis on technical aspects in a data warehouse project makes it hard to change any type of business process and especially core processes. It is often the case that organizational aspects of data warehouse development have a low priority, and this may result in that actors in the organization do not realize the benefits of changing the core processes. By neglecting the organizational aspects of the data warehouse development, the problems with ownership and low degree of participation will arise. By paying more attention to organizational aspects, it may be easier to involve actors in the development and thereby the problems may be reduced and it will be easier to suggest changes also in the core processes.

Supporting processes is one type of business process that is frequently mentioned as often being exposed to needs for change. Supporting processes are not as central in the organization as are the core processes, which make them easier to change without directly affecting the development of products and services. Another reason is that this type of business processes are often less optimized and documented, which makes the need to change them more obvious to the actors.

Examples of supporting processes affected by the introduction of a data warehouse are e.g. processes that concern the relation with customers and suppliers. Data from these processes is a strong candidate to be used by the data warehouse and is thereby crucial, with respect to error free and high quality data, in order to create a data warehouse with acceptable performance.

Management processes are also considered easier to change than core processes. The reasons why management processes are considered easier to change than core processes are the same as for supporting processes, i.e. low degree of documentation and optimization.
Management processes concern strategic and tactical planning, which are supported by management information systems. As stated in chapter 6.1.2: it is obvious that the introduction of a data warehouse project gives rise to a need for change. It also adds some new features to the way work is performed at the higher levels when a data warehouse project is introduced, since the actual work at strategic and tactical levels will be changed by the introduction of a data warehouse. The reason for this is that a lot of work related to strategies and tactics will be performed by using the data warehouse instead of the existing systems and techniques used to gather strategic and tactical information.

### 7.2 Objective 2 & 3

There are many factors that may trigger a need to change business processes and a majority of these triggering factors are not specific for data warehouse development, but are general factors in context of ISE. The introduction of an information system in an organization works as a catalyst, which always results in a need to change how some things are done. Nilsson (2000) and Cherry & Macredie (1999) points out the relationship between the reality and the information system. Both reality and information system affects each other and changes at one end will require changes at the other end.

As stated in the analysis, the triggering factors that cause a need for changing business processes may be divided into different levels of abstraction: general level, type level and instance level. Triggering factors at a general level are:

- The achievement of maximum ROI may result in a need to change business processes. An example of this is the production of required data in the source systems. Without error free and reliable data, the data warehouse will not perform satisfactory. This triggering factor may be considered as general for any information systems development but in this context it is different since data warehouse development projects in many cases are more costly, which results in a greater focus on ROI. Another distinguishing feature is that a data warehouse uses a bigger amount of data than other types of information systems.
7 Result

- The data warehouse project itself might trigger a need for changing business processes. The introduction of a data warehouse project makes it possible for actors to view certain aspects in a new way, which might result in new demands to how operations should be performed. The distinction from other types of information systems development is that a data warehouse project controls the data more rigorously and the focus on data and definitions is greater.
- New organizational knowledge might evolve, since the data warehouse project requires analysis of the organization in order to state data warehouse requirements. By doing this, actors in the organization might be forced to reconsider how things are executed at present state and what business processes that need change in order to reach a desired future state. A data warehouse project often covers many or all different parts of an organization, which results in a more extensive analysis of the organization, which in turn results in a better overall picture of how things are done, than other types of ISE projects.
- An organization will probably have many different goals with its intended data warehouse. All goals will not be achieved, since this would require too much resources and time. In order to get maximum ROI the organization is forced to prioritize its goals in order to select those that are considered as the most important. By being forced to prioritize, organizations will pay more attention to related business processes in order to be sure that the most important goals are selected. The organization must know what it takes, in context of business process change, to go from present to future state.

At a types level, some correspondences between triggering factors and affected types of business processes have been identified. As already stated, all types of business processes are affected by a need for change during the pre-deployment stage of a data warehouse project. We have chosen not to focus on supporting processes and management processes, since these are considered easier to change. Although, many of the correspondences presented may be applicable to core processes as well.

Triggering factors that creates a need to change supporting processes are:
7 Result

- The fundamental reason to create a data warehouse, i.e. better knowledge about customers and suppliers that creates a will in the organization to improve way of working with concerned business processes, such as logistic processes.
- The need to produce the required data in the source systems. Data from logistic processes are one type of many that is used for this.
- The need to produce error free and trustable data in the source systems, which will result in maximum ROI.

Triggering factors that creates a need to change management processes are:

- Increased competition, which requires better and faster decisions about marketing etc. This is a fundamental reason to create a data warehouse.
- Proper alignment between business processes and data warehouse will increase the opportunity to find new customer segments and market shares. It is important to have in mind the strong relationship between the organization and the information system, which affects one another.
- By being forced to prioritize its goals, the organization must consider its business vision, which might result in a new way of viewing the business processes concerned.

At an instance level, we have not found any general correspondence between certain business processes and the triggering factors. Material from interviews indicates that the correspondence between business processes affected and the triggering factors are mainly project specific.

With respect to the aim of this thesis, we have investigated which factors that may cause a need for changes in business processes during the pre-deployment stage of a data warehouse project. The factors have been identified in chapter 6.2 and chapter 7.2. We have also investigated and presented (in chapter 6.1 and 7.1) business processes that are more
7 Result

commonly affected and also how these business processes are affected when a data warehouse project is introduced.

From the above, we consider the aim and objectives for this thesis reached.
8 Discussion

This chapter reflects on the conclusions and results found in this thesis. In addition, suggestions for future work are presented and discussed.

8.1 Reflections on results

In the introduction of this thesis, the relationship between data warehousing and business processes is discussed. It is stated that a data warehouse is of vital importance for decision makers and the conclusion is made that the relationship between data warehousing and business processes may be used as an opportunity to change business processes in an organization. We have paid attention to the pre-deployment stage of a data warehouse project, since changes of business processes at this stage, will prevent changes that have to be done later on, probably at a higher cost.

In order to reach the aim and objectives, literature studies and interviews have been performed. The gathering of literature has not encountered any problems but the interview study has run into a few problems. First of all, the contacts with the appropriate respondents have consumed a lot of time, but after consulting a data warehouse consultant (Lars Boström) with the appropriate connections, suitable respondents were identified. Another problem related to the interview study has been the availability of the respondents. We had preferred if the interviews had been made in a face – to – face manner, but this was not possible, since the respondents did not have the time available.

The results from this thesis clearly indicate that there is a need to change all types of business processes at the pre-deployment stage of a data warehouse project. Although, the core processes are considered hard to change at this stage because of their important role in the organization. The post-deployment stage of a data warehouse project, i.e. the stage that concerns maintenance and management of the data warehouse, might be the stage where these core processes may be changed, since data stemming from the source systems, based on core processes such as manufacturing and sales, might produce results that are clearly incorrect. By convincing the actors in the organization that the core processes produce incorrect data, a change of these is possible, but will probably consume a lot more
Discussion

resources since the data warehouse is up and running at this point in time. Other triggering factors concerning the post-deployment stage is new knowledge derived from the data warehouse, which might indicate that some business processes may be performed in other ways than present, which improve their output. An aspect worth mentioning is that the identified triggering factors related to the pre-deployment stage of a data warehouse project have a lot in common with factors present in other types of ISE projects. This is not a surprise, since a data warehouse is also considered to be an information system, which is used in its own characteristic way. Although, the triggering factors related to data warehousing differ a bit compared to “ordinary” information systems related factors. The main characteristics that make a difference are for example that triggering factors related to data warehousing pay more attention to ROI, because of the high costs associated with data warehouse development. It is also a greater focus on data since the amounts of data used is very large.

One problem that has occurred is the separation of the different tracks in the business dimension lifecycle diagram, developed by Kimball et al. (1998), described in chapter 2.2. The diagram describes a typical development cycle of a data warehouse, and uses four different tracks that handle different aspects of the development (data, technology, application and management). The problem is that no track is assigned to deal with business processes and it is hard to assign a certain business process to just one track, since it is often the case that a business process spans over more than one track. This might result in the “business process stove – pipe syndrome”, i.e. the same business process is analyzed at all four tracks at the same time without any synchronization. This is a weakness that is not highlighted enough by Kimball et al. (1998). By deploying a greater business process focus in the business dimension lifecycle diagram, this syndrome might be reduced. The characteristics of a business process, mentioned by Willoch (1995) in chapter 3.2, is another problem that may affect the handling of these. It is mentioned that business processes do not have clear names, are invisible and do not have any one responsible for a certain business process. Business processes are by nature a bit tricky to handle, since they are hard to make clear and explicit. These characteristics will not make the assigning of
8 Discussion

business processes to certain tracks any easier, which stresses the importance on paying more attention to business processes when developing a data warehouse.

One other interesting subject is the development of data marts, which consumes fewer resources, i.e. time and money, than the development of a fully-fledged data warehouse. By creating data marts that relate directly to related business processes, such as sales and advertising, it might be possible to control and change related business processes easier. If the data marts were built properly, a future possible integration of these into a data warehouse would not be an impossible task.

Another issue worth mentioning is the number of respondents used in the interview study. It might be considered that eight respondents is a small number, but it is important to consider some factors that have strongly reduced the number of respondents. Firstly, the interviews have been made with respondents employed in Sweden, which decrease the number of possible respondents. Secondly, there does not exist a great number of companies specialized in data warehousing located in Sweden. Thirdly, since we have been focusing on the organizational aspects of data warehousing as well as the technical aspects, the number of possible respondents available is small. Finally, not all respondents asked to participate in the interview study have been interested in participating. Together, these factors have reduced the number of respondents, but we consider the participators experience and knowledge within the research area as a contribution and the answers given have made clear indications and a lot of information with respect to the aim and objectives for this thesis.

The main contributions from this thesis, i.e. a characterization of factors that may affect business processes during the pre-deployment stage of a data warehouse project need to be further investigated in order to give the ideas presented a more solid ground. However, the ideas presented should make the building of data warehouses a bit easier since we have highlighted the need to change business processes. We have also shown reasons to why it might be hard to introduce changes in business processes to actors within the organization. We have also presented some suggestions how to solve these problems, or at least reduce
them, which probably will make the introduction of changes easier. The work will also be a good starting point for future work, which will make the ideas presented in this thesis more solid and probably add new ideas and features that will decrease the complexity and the resources used, when building data warehouses.

### 8.2 Future work

We have identified three ideas, which would be appropriate in order to follow up the results presented here. We intend to describe each of these ideas, and motivate why these ideas are important.

#### 8.2.1 Evaluation of results and conclusions

It would be interesting to investigate the results from this thesis and evaluate these in a data warehouse development project. If the outcome from a project study confirms the ideas presented in this thesis, the results presented may be considered to be more reliable. It would also be very interesting to evaluate if the correspondences presented in this thesis at a type level, also are applicable at an instance level. If the opposite occurs, necessary proposals for change may be easily identified. In any case, independent of positive or negative result, the results and conclusions presented in this thesis will stand on a more solid ground.

#### 8.2.2 Replication with new respondents

Another interesting idea is to use the same aim and objectives used in this thesis, and apply the same questions, but on new respondents. By doing this, the results and conclusions presented in this thesis would be evaluated, and hopefully verified. If the opposite occurs, the reasons for this would be interesting to investigate and the new results would be more reliable. An important factor for the possibility of replication is the selection of respondents. It is important that the respondents not only has a technical background within data warehousing, since the organizational aspects have been highlighted in this thesis, i.e. the respondents must have the technical knowledge behind data warehousing but also have a good understanding of the organizational aspects of data warehousing.
8 Discussion

8.2.3 Study the difference between failure and success
It would also be interesting to investigate if there are any distinguishing features between data warehouse projects that have been successful and those that have not been successful. First of all, it might be interesting to verify and perhaps contra state general reasons for failure already identified in literature. Another interesting part, with respect to the focus of this thesis, would be to investigate if there are any differences with respect to changed business processes. If there are any distinguishable differences between successful and unsuccessful data warehouse projects, the results from this thesis will be even more important to investigate further.
References


References


References


References


Appendix 1

Appendix 1: Questions to respondent.

1. Name?
2. Company?
3. How many years have you worked with data warehousing?
4. How many different data warehouse projects have you participated in?
5. What roles have you had in these projects?
6. What is your description of a data warehouse?
7. What is your description of a business process?
8. Is it common that requirements on data warehouse results in a need to change business processes in an organization?
9. What factors may be a trigger for such need?
10. Is it any factor that affects business processes more than other?
11. What type of business processes is most often affected by a need for change with respect to data warehouse development?
12. Is it possible for you to categorize the given processes into core processes, supporting processes and management processes?
13. Is there any difference, with respect to affected business processes, between successful and not successful data warehouse projects?
14. Do you see any general correspondence between any triggering factors and business processes that give rise to a need for change when developing a data warehouse?
Appendix 2: Answers from respondents.

Respondent 1:

1. Örjan Lövström.

2. Guide.

3. I started with warehousing 1994 and has been involved in it since then. I worked a lot with web based systems a while ago but now I am back with warehousing again.

4. I have been involved in ten projects.

5. My roles have been advisor in combination with project management, especially at the analysis phase of the project.

6. I do not see DW as a database, but more like a concept that stores information that is important for the company. DW also stores in which way the information is important as well as common concepts for things in the company. What is want is one version of the truth and this requires that everybody have an agreement of definitions of concepts. I also see a DW like you do in your report. (Chaudhuri & Dayal, 1997)

7. When I hear the concept business process I refer to a typical core process, i.e. customer and manufacturing processes.

8. Yes. This need for change is necessary in order to get out the full potential of the DW. In order to get this out you have to change your way of working.

9. The adjustment of way of working which is necessary if you want to get out the full potential from the DW.

10. I cannot respond to this question.

11. Those processes that are supported by the DW must be changed which in turn generates an increased awareness of the business processes in the company. DW works like a catalyst. It is also important to remember that those problems that are identified during the project shall be handled in separate projects, otherwise the DW project will be suffering.

12. Processes that produce garbage data in the source systems is mostly related to supporting processes.

13. It is common that an unsuccessful project is collecting a lot of data that actors want to involve in the DW. The problem is that sometimes the actors do not know what
they shall use the data to. This might be a big problem that slows down the project and make it less efficient.

14. I cannot respond to this question.
Appendix 2

Respondent 2:

1. Torbjörn Leise.

2. Unisys.

3. I have been working with DW for two years now.

4. I am currently working in my first DW project.

5. Developer and manager.

6. I view a DW like a system for collecting data from different sources that is compiled in a report tool.

7. A business processes is the way that work is performed.

8. My experience is that the DW has been used as a report tool for follow up. Business processes has been changed after this follow up, but at this time the DW was already up and running.

9. The need to follow up work at the company triggered the need to create a DW.

10. I have no answer to that.

11. I have no answer to that.


13. I have no answer to that.

14. I have no answer to that.
Appendix 2

Respondent 3:

1. Pehr Buhrenius

2. Tridion AB.

3. I have been working with warehousing for 20 years in different forms.

4. Approximately 20, but I am not sure. The last really big DW project I been involved in was at Ericsson there we built a business warehouse and showbiz production. Many projects have been smaller and have a lot of common with data marts.

5. Project manager and advisor for users in the company. My roles have often been at a high level and angled from the user side, i.e. helping users to see possibilities and needs.

6. I see a DW as clean, washed data that is stored in one single place that is being fed from several transaction systems. The DW consists of both Meta data and master data. Data in a DW shall be used for analyses and to get out data and it is really important that the out data is not incorrect.

7. I view a business process like flows. Something that takes an input and increases it until it reaches output. A business process is a value increasing process that promotes a business transaction.

8. Yes. It is important to define a certain scope for the changes and it is critical to show the benefit with a certain change for the users in the company. An ordinary way to do this is to first look at the present situation and then look at the future situation that we want to reach. After that it is investigated how much it will cost and what changes that needs to be done. If you want to get better you have to change your processes. Simplify and improve is the motto.

9. The triggering factors are mainly the definition of values and the efficiency that is demanded in order to get from present to desired future state. It is critical to be able to show/convince the user the benefit with a certain change in order to get the DW to work in a better way. It is important to get rid of suspicions in a company and kill the myth that all processes are holy.

10. Hard to say.

11. The usual is that you make changes in the organization in different ways. You want to make things easier and question the way things are done at present time. Everything shall relate to a value, if it does not, throw the process in the garbage can.
Appendix 2

12. Mostly it is management processes and sometimes supporting processes. With a DW you want to increase the customer beneficence and this in the past this often resulted in slimmed production and related supporting processes. Although, I think that a core process cannot be improved by itself; you have to start the improvement at the management level, since everything in the organization is directed to sell a certain product to a customer. If you do not have any customers, it does not matter that your core processes are excellent.

13. Hard to say.

Appendix 2

Respondent 4:

1. Magnus Tegerup.

2. SYSteam AB (Coordinator for selling and market for the cooperation). I have previously been responsible for market and selling of BI/DW at Guide Consultant. Previous work as management consultant and VD. Founder of the network “Data Warehouse Group”, that is a constellation of a number of companies the focuses on decision support systems.

3. 4 years.

4. 10.

5. Seller, consultant with focus on business developing, responsible for orders.

6. There are different definitions that and most of them are correct in their sense. I want to describe a DW like: One ore more databases, there the organization stores/collect important information for operative, tactical and strategic decisions.

7. A business process is according to me those processes or flows that stream through an organization and is related to a direct customer interest. Refers to primary activities in M. Porter value chain.

8. Yes, this is almost always the case. Most of all, the users of information gets aware of that they have to prioritize what information that is really needed, i.e. it is often the case that the customer wants it all and that is not realistic. It is when the project management gets weak that DW projects as taking to long and consumes too many resources.

9. New products or services. The degree of service. Competition. Requirements from suppliers and customers. Fusions with other organizations. Expansion or reduction of the organization.

10. Hard to say but I think that requirements from suppliers and customers is the factor that affects the management of the organization most.

11. Something that is often discussed is logistics with suppliers. Other things are product development and advertising campaigns (Time to market). In the service sector is it the satisfaction of customers that makes the difference between competitors. Companies with stated concepts about customers are more probable to be successful at the market.

12. Everything is related to the core process, which is the most obvious when concerning organizational success. Customer satisfaction (supporting processes) is
only one part of the chain that consists of different activities where relation and
degree of service from our suppliers affects our customers.

13. If we speak about pure DW projects, the most important question is how big grip
the company wants to take, i.e. if you as a company try to change everything at the
same time it is most probable that you will not get any success. Because of this, it is
important to make prioritizations in the project. This is important since the
development will not take too long time. Too long development time renders in
poor motivation and interest about the project from the actors in the organization.

It is probable that exists general truths but often, triggering factors are unique for a
certain project.
Appendix 2

Respondent 5:

2. Ecsoft.
3. 5 years.
4. 5 projects with various sizes.
5. In the beginning my roles was database developer/DBA and programming a lot of OLAP – cubes etc. Recently, I have been working more with the organization, development of requirements specifications. I know both sides of the coin, i.e. technical and organizational aspects of DW.
6. Fun question, I use to say that a DW is an architecture that is kind of technical, i.e. like the framework presented by Chaudhuri & Dayal (1997). I think that DW is an architectural principle.
7. I think that a business process is something where you put something in at one end and in the other end something more valuable comes out. Typical business processes are selling processes and customization processes. Although I believe that the customer is not a central in the business process as it used to be, with respect to customization processes. A business process is a process that generates a value to both company and customer.
8. Yes it is common. It can be the case that you have to change business processes and that you have to gain more knowledge. One important thing is that actors that are working within the business process understand how the DW may improve their work. This is something that often is skipped, which results in less work done with changing processes. The biggest problem with change is that it is too many technical consultants involved in organizational aspects; there it should be more management consultants instead.
9. The triggering factors are dependent on what the DW intention with the DW is. If you want to collect information and reach increased knowledge and it is the processes that are concerned, then this is precisely one factor with respect to previous question. If we then have gained that increased knowledge it is probable that we can affect the way the processes are working. By knowing more about the process we can make big or small changes to it.
10. Most of all knowledge about the processes.
11. Most of all the selling process that is affected directly by changes. Also CRM but most of all change there you can look at what products the customers want, where they are and what selling channels we shall use in order to reach these customers.
Appendix 2

Also manufacturing processes but these are heavy and hard to change. The selling process is the most usual and easiest to change.

12. Supporting processes and management processes are those business processes that is most exposed the change. The core processes are too hard to change. I think that core processes also should be affected but these processes have been used for long time in the organization and have been improved over a long time.

13. If you do not focus enough of the actors in the organization that are going to use the DW and be participating in the development and pays more attention to the technique instead you will get into trouble. It is important that you have both cultural implementators and technical implementators. It is often the case that the work with change is lost when the technical aspects gets too much attention, which often is the case when the consultant side is to technique oriented. One other thing is that it is important to have support from the company’s management board when introducing a project.

14. With respect to the need for change is it that you do not have control of which processes that is profitable or not. It is also common that you do not have the knowledge to use all the data that exists in the company. It is important to have the ability to change your way of working if you want to reach new information.
Appendix 2

Respondent 6:

1. Fredrik Hjelm.

2. Posten IT.

3. I have been working with warehousing for 4 years. During these 4 years I have been working mainly with one big project but also with one smaller project. The smaller project was a temporary DW, which made the building of the big DW a bit easier. Before I started working at my present employment, I have been working with other systems that looks a lot like warehouses but with less features.

4. Three projects that concern pure data warehousing.

5. In the smaller projects I have been working as a project manager. In the bigger project I have worked as in data responsible in parallel with responsible for quality. Later on, I have worked as a project manager in the big project as well.

6. The best definitions I have found are mentioned by Fredrik Söderström. The first definition I would like to mention is: A DW is a logically integrated data set, which is intended for analysis and mirrors the company by collecting data in registers regularly. I think that this definition is a bit formal and technical. Because of this I have selected another definition: You may see the DW as a structured storage place, a grocery, where the starter is placed at one location and the dessert at another location. The grocery isn’t structured by suppliers, but is structured in order to be easily found by the consumers. In the DW, the storage has been adjusted by a certain need in order to be easily found by the users. I like this definition better. Although, it is hard to say what a pure DW is since circumstances and views is very different in different companies. Although, it is important to not violate any of the fundamental rules, i.e. non-volatile and historical data. It is important to always change the data in the source and not in the DW.

7. Something that generates an increased value for the company and is able to get paid for.

8. Both yes and no. When you are collecting requirements it is most likely that you experience a need to change some business processes. But actually change theses processes is nothing that you can do very easy, especially in a big organization. It may come up some indications on a need for change but these needs are more obvious once the DW is up and running. At this time, it is very common that strange things are observed, and the users of the system realize that some changes are needed. Big shortcomings in quality are very common to detect at this stage and it might take a little time before the users realize and accept this. Often, it is quite hard to say that different departments systems in an organization don’t fulfill the stated requirements, but after a little time has passed, these problems are easier to
introduce. My experience is that it is very hard to introduce actual changes before the users themselves have realized these.

9. Often, new ideas about possibilities are generated during the project, which may result in a need for change. Another factor that might trigger a need for change is to increase that quality of the data that is going to be used.

10. Hard to say, I think that it depends on the definition of changing a process and how big the change is; often there are other things than a DW that affects the change of a process.

11. Supporting processes are common, if you consider that quality aspects. I also consider all processes that concern customers in any way as highly likely to experience a need for change.

12. All three types of business processes mentioned in your work are considered, but with a greater focus on supporting processes.

13. It is important to have the managements support. Another thing is to focus on solving questions and not only generate new questions without solving.

Appendix 2

Respondent 7:

1. Jonas Holmgren.

2. Component Software Sverige AB.

3. 2 years.

4. I have been participating in the selling process in a number of projects.

5. Participating in the selling process.

6.

   a. A data warehouse is time stamped base data collected from the organization.
   b. A data warehouse function is to make sure the decision makers in the organization have access to correct information and has the possibility to trace all aggregated and summarized information to its original source.

7. A business process is a process that is executed of a company’s co workers in order to fulfill a service. An example of business processes that is present in any organization is delivery processes, purchase processes and selling processes.

8. No, that is not a requirement. That requirement that is often mentioned during a project is that a data warehouse shall make present processes more efficient and make it possible to make better decisions. In those cases where the business process shall be changed is it often the case that the definition of a new process requires a data warehouse in order to make the process real.

9.

   a. Competition, competitors have a more efficient process.
   b. A desire to decrease costs.
   c. Changed/more efficient organization.
   d. Increased volumes.

10. Dependent of business cycle. In an economic upswing is it increased volumes/sales and in economic regression is it cost savings.

11. I am not sure but the type of business process that often is in focus in the context of making processes more efficient is processes that is depending on a number of source systems and transaction intense processes.

12. This type of processes mentioned above is supporting processes and core processes.
Appendix 2

13. Yes. The projects that have been successful have reached good payback and have changed the processes in a positive way. The less successful projects often have problems with delivery contracts (data quality) and long loading times.

14. No, everything is project specific.
Appendix 2

Respondent 8:

1. Jörgen Andersson.

2. Oracle Consulting.

3. 4 years.

4. 9 DW projects.

5. Project manager (PM), system design (modeling) and systems engineering.

6. There is authors to many books within the area that has many good definitions, all definitions does not fit in all situations and as a consultant you have to adjust the definition to the customers needs. Thereby it is the customer needs that decide what definition that is appropriate for a certain project. Personally, I use to say that I am working with the creation of information from data. This description is usable in any situation. Oracles official definition of a data warehouse is: “An enterprise structured repository of subject-oriented, time-variant, historical data used for information retrieval and decision support. The data warehouse stores atomic and summary data. This definition stems from Oracle Data Warehouse Method.

7. Structured, documented/defined activities within a company that creates value for the company, direct or indirect. (Making money). Precisely as you divide the processes in types (core, support and management), are there also different levels of processes. One high-level process may consist of several sub-processes. This is of course deflecting the design of the data warehouse.

8. I can identify three cases which lead to changed processes:

   a. Information in the data warehouse reveals bottlenecks in the processes. By making reactive measures, the company can change processes in order to avoid bottlenecks.

   b. By using the information in the data warehouse the company may not only just analyze performance in its processes, but may also perform “what if” scenarios and thereby find out new ways of increasing performance in some processes. This does not always lead to changed processes, often the result id that the company pays more attention to some more profit making activities, i.e. the company chose to go for one advertising campaign but not some other campaign. (Proactive measures)

   c. The data warehouse is designed according to the company’s documented business processes. When data is loaded, and finally analyzed, it is shown that the documented processes do not agree with how the company actually works. The gives the result that measure points in the data warehouse do not contain any valuable data and leads to that processes can not be analyzed in
Appendix 2

a good way. In this case it has to be decided if processes are incorrect documented or if the processes are carried out in the wrong way.

Case a and b is discovered at the pre – deployment stage when the data warehouse is already developed and is up and running. Changes at this stage demands measures that concern management.

Case c is more complicated to detect and is important to detect as early as possible. During the past two years, Oracle has worked hard with changing its Data Warehouse Method in order to detect these problems as early as possible. The method is based on Time – boxing and prototyping, in order to reveal discrepancies between the reality (in data) and the, by the company described reality. By using prototypes where real data is used, these changes can be found early in the project.

9.  
   a. How well the company knows its processes.  
   b. Which key persons in the company that is available in the projects definition and design phase.  
   c. Access to the right people that knows the data in the source systems.  
   d. Those persons that is assigned to design/develop the system. By asking the right questions from different angles (cross questions), discrepancy can be discovered before you start building a prototype.

10. Hard to tell, all factors seem to be equally important.

11. Hard to say. In all projects I have been involved in I feel like it has been equally divided. By intuition I think that supporting processes is worse documented and should thereby fall into the frame of change. Core processes is often well documented and the personnel knows these processes very well.

12. See above.

13. The difference is how the project is carried out. A well planned project with the right actors and the right method may be successful even if the company has poorly documented processes.

14. I cannot see any general correspondence.