Assessing Business Intelligence Business Value in Decision Support Systems

A Study at a Nuclear Power Plant

GÖRAN LINDMARK

KTH Electrical Engineering

Master Thesis
Stockholm, Sweden 2007

XR-EE-ICS 2007:04

Abstract

Since the accident at the nuclear power plant of Chernobyl in 1986 the work with international industry information exchange, benchmarking and best practice sharing, has accelerated. Assessing the performance of the nuclear production processes is a key issue when working with the improvement of safety and production levels globally. Process controlling and effective and accurate internal information sharing is also a competitive advantage for every power generating company. To take better decisions, based on process assessment data, could be made possible in different information system scenarios. To reach these possible future systems scenarios would require a certain amount of effort in modifying the existing decision support systems.

In this master thesis the non-financial processes, sub-processes and key performance indicators from the international process performance assessment standard, the *SNPM reversion 3*, have been used to measure the process assessment degree at one nuclear power plant of a utility company. A cost/benefit analysis, where cost is measured as modification effort and benefit as process assessment degree, has been performed to decide which of three future system scenarios that would provide the highest Business Intelligence business value.

The study shows that the low modification effort of one scenario is compensated by the high process assessment degree of another scenario and the result implies that none of the proposed system scenarios is proved better than the other. Therefore no scenario recommendation can be made based on the factors assessed in this study.

Keywords: Business Intelligence, Decision Support Systems, Data Warehouse, Data Mart, Process Assessment, Modifiability, Extended Influence Diagrams, Scenario Evaluation, Standard Nuclear Performance Model.

Table of Contents

1 Nomenclature

- BI – Business Intelligence
- CFU – Chef FöretagsUtveckling
- CSF – Critical Success Factors
- DW – Data Warhouse
- DM – Data Mart
- DSS – Decision Support System
- EID – Extended Influence Diagram
- ICS – Industrial Information and Control Systems
- KPI – Key Performance Indicator
- KTH – Royal Institute of Technology
- ME – Modification Effort
- MIS – Management Information System
- NEI – Nuclear Energy Institute
- PAD – Process Assessment Degree
- R&D – Research and Development
- s0 – Scenario 0
- s1 – Scenario 1
- s2 – Scenario 2
- s3 – Scenario 3
- SNPM – The Standard Nuclear Performance Model
- WANO – World Association of Nuclear Operations

2 Introduction

This Master Thesis aims to describe the methods and the theory that has been applied in a case study at the utility company's nuclear plant. It also describes how the empiric data has been collected and how the analysis of the data has been performed. Finally the result of this study is presented and a discussion about the study in general and the result in particular is held. First, though, a short background for study is given.

2.1 Background

The utility company has for more than ten years been working with Business Intelligence questions and how data from different operations could be made available in a greater extent to the whole organisation [55]. In organisations with a very federal management structure, such as the utility company, the different departments and operation fields tend to attend to their own business only and information exchange is generally very low [38]. One field where enhanced information exchange would be particularly interesting would be nuclear operations [27].

Before 1986 nuclear plants were operated on a site basis where experiences and best practices were in best cases kept on a national basis. Since the accident at the Chernobyl nuclear power plant, nuclear operators world-wide realised the dramatic consequences accidents could have globally and a need for international cooperation arose [52]. To know the operations in other plants do not only give a safer sensation but also makes it possible to learn from each others mistakes and follow best practices. The World Association of Nuclear Operations (WANO) was formed in May 1989 by nuclear operators world-wide uniting to exchange operating experience so that its members can work together to achieve higher standards of nuclear safety and other operational goals [52].

The knowledge field of Business Intelligence is just as the organisation WANO something which has grown in importance during the last decade of the last century. What was initially a much undefined area of knowledge about how to handle the ever growing availability of operational data and turn it into useful information, has been refined [18]. The many trail and error Business Intelligence projects makes us today aware of what techniques actually works, which that doesn't [22]. One of the great challenges is how to unite data from a whole organisation in a common database in a way that new analysing techniques can be used in order to derive information about cross organisational trends and dependencies [28].

One of the core aspects of Business Intelligence is the use of Key Performance Indicators. They can be used both for benchmarking one business with another and as goals to work towards. If benchmarking should be possible different businesses needs to define Key Performance Indicators in the same way. The Nuclear Energy Institute (NEI) have, with the support of WANO and other nuclear branch organisations, begun the work towards a standard how to identify nuclear power plant processes and how to define Key Performance Indicators. [39]

To conform the Business Intelligences systems, of an organisation, towards international standards, great modification work might be necessary. A huge challenge is how to measure the effort needed to implement such modifications and how to evaluate the value of certain change. The Department of Industrial Information and Control Systems at the Royal Institute of Technology perform research in both fields. [21]

2.2 Problem

The problem that the utility company faces is how to shape the Business Intelligence systems, so that they support global standards as well as reassure that the data in the company transform into information that can be used as competitive advantage.

2.2.1 Problem relevancy

Knowledge and information have always been regarded a key competitive advantage in all businesses. As the amount of data grows larger (often over 20 Tb in large companies) it is every day more important to find solutions, which can provide full range functionality and are trustable and won't fail in execution. [12]

Benchmarking has proved being the best performance increaser in the nuclear industry. The ongoing work to define common processes and performance indicators is the basis for all benchmarking, since comparison is useless unless the same things are compared. [39]

2.3 Question at hand

The study aims to answer the following question:

Which future system scenario enables the highest Business Intelligence (BI) business value in the Decision Support Systems of the nuclear power plant?

Considering the delimitations and assumptions made, the following clarifications might be necessary to understand the question:

- Only the three predefined scenarios described in this report are considered, see further in *2.6.4 Delimitations in scenarios* .
- BI business value is measured in Process Assessment Degree (PAD) and Modification Effort (ME):
- PAD is measured in the number of available SNPM KPIs, see further in *2.7 Assumptions*.
- ME is measured in terms of effort needed to do a certain change on a certain set of systems.

Sub-questions:

Given the main question above the following sub-question have been identified in order to support the main question:

- How can the BI business value be assessed?

2.3.1 Reason for the question

According to Yin a case study like this one is proper for answering questions that are descriptive, diagnostic or normative [54], provided that Gammelgårds definition of questions are used [15].

Question	Example of question
Investigative	What is happening?
Descriptive	What does it look like?
Explaining	Which connections exist?
Diagnostic	Why did it become this way?
Evaluating	What did it result in?
Normative	How should you handle this?

Table 1, Question types and question examples.

The main question asked is clearly normative since it is going to answer in which way one should handle a situation.

2.4 Goals

Goals can be of two kinds: Project goals and Business goals. [11]

The following Project goals have been set for this study:

1. Define the concept of "Business Intelligence" and other important concepts associated with it, so that their function becomes clear.
2. Develop frameworks for measuring both the PAD and the ME.
3. Elicitation of three future scenarios between which the PAD and the ME can be compared.
4. Perform an empiric study in one of the divisions at the utility company that leads to enough data to carry out the analysis.
5. Carry out the analysis in a way that it becomes clear on which basis a future scenario could be recommended.
6. Make presentations of the study in forms of oral presentations for both the utility company and KTH, a written report and an executive summary, which clearly shows on what basis the study has been carried out and which result it has reached.
7. To deliver a substantial value to both the utility company and KTH throughout the lifespan of the project.

The following Business goals have been set for this study:

1. The project should all in all take no more than 800 man-hours.
2. The project should be completed and fully presented before the first of February 2007.

2.5 Conditions for the study

Time is the greatest restraint of this study. Since it is carried out as a master thesis only about 800 man-hours should be spent.

The second condition, that is important to mention, is the fact that nobody at the utility company, except of the mentor for the study, has been given resources to set aside time for this study which in practice means that all cooperation with people at the utility company has been done on their "spare time". And people's spare time has been very limited during the time for this study.

2.6 Delimitations

This chapter describes what delimitations have been set up due to the study's conditions or other reasons.

2.6.1 Delimitations in objects included in the study

Within this study only the Business Intelligence systems of the nuclear power plant will be studied. This is due both to the time limitations of this study and the availability of study objects. The study is, however, performed in such a way that future studies, on other study objects, can be done on the same basis as this one.

2.6.2 Delimitations in types of systems

This study restricts itself only to look upon Business Intelligence systems, which provide decision support of a technical kind. That means that all systems only serving the processes related to finance, accountancy, and human resources, will be excluded.

2.6.3 Delimitations in types of KPIs

Only “technical” indicators connected to the processes of generating power have been included.

2.6.4 Delimitations in scenarios

Only three different future scenarios will be considered. There are theoretically an unlimited number of possible future scenarios but the three elected have by the mentors at both KTH and the utility company been considered the most interesting and likely ones.

2.7 Assumptions

There are a few basic assumptions for this study.

- Business Intelligence business value can be measured in the terms Process Assessment Degree (PAD) and Modification Effort (ME).
- PAD is measured in number available KPIs.
- The only KPIs of interest to measure are the ones provided in the SNPM; see further in *3.4.2 Building the PAD framework*.
- ME is the work effort to do a certain change on a certain set of systems.
- The study object is assumed to have more than one system with Business Intelligence functionality and not to use any Data Mart or Data Warehouse.

2.8 Disposition

This thesis has the following disposition:

1.	Nomenclature	Abbreviations often used in the report are explained.
2.	Introduction	This chapter gives the general base for study. The question at hand, the delimitations and the assumptions made.
3.	Method	The method is described first on a general level then more specifically for the building of theoretical frameworks, the empiric collection of data and the analysis.
4.	Theory	The theory has three main parts. The background information about especially Business Intelligence helps forming the second part of the theory; the scenarios. The third part is the theoretical frameworks that are the basis for all empiric research and the analysis.
5.	Empirics	This short chapter describes the empiric data collection. The data itself can be found in the Appendixes.
6.	Analysis	Here a detailed explanation on how the results have been calculated is shown.
7.	Results	In this chapter the result of the study is presented and the question for the study is answered.
8.	Discussion	In the discussion chapter the result itself is discussed and the factors leading to the result. Strengths and weakness with the study in general is also discussed. Finally examples on improvement potential areas for the nuclear power plant are presented.
9.	Future work	Suggestions for further studies are presented in this chapter.
10.	Acknowledge	Contributors to the study is thanked
11.	References	The references for the report are found here. Note that references shows in the appendixes are presented in the end of each appendix.

3 Method

The purpose of this chapter is to give a well described picture of how this study is carried out, which general phases and elements this project consist of. It will also give an insight in of what information has contributed to build the study's frameworks. Furthermore a description of how data have been collected, on which bases they have been compared and the method used for the analysis, is presented.

3.1 The general study process

The study can be divided into five parts: Project establishment, Theoretical study & Problem approach development, Empirics, Analysis and Presentation & Project closing. The general idea is that these parts follow each other sequentially but in fact, during the process, work with different parts continues parallelly.

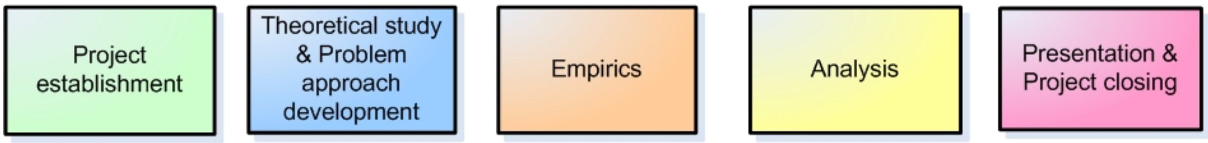


Figure 1, the study process.

During the process a number of scenarios for the Business Intelligence systems at a nuclear plant at the utility company are studied with the aim to determine which scenario gives the most PAD, measured in available KPIs, and how much ME each scenario requires.

In the following sections the different phases are described more in detail:

3.1.1 Project establishment

The project establishment is the first part of the study and consists mainly of creating good contacts with the interested parties and specifying the basic guidelines, according to which the project is to be carried out. Initially all parties sign an agreement of terms for the project and later an initial report is created.

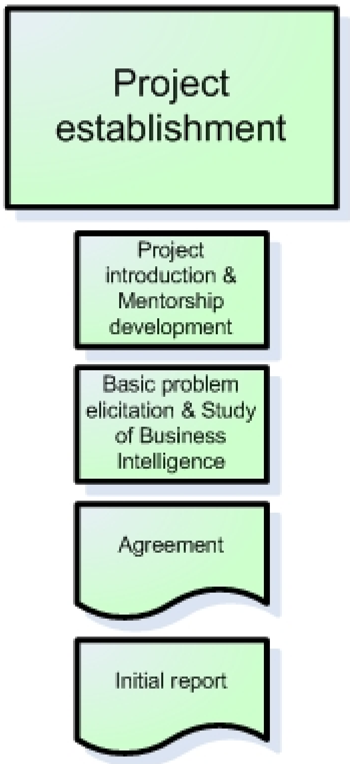


Figure 2, Project establishment and its elements.

3.1.2 Theoretical study and problem approach

The second and very important phase includes a massive theoretical study and further development of the problem, setting up delimitations and selecting units that will be included in the study.

The general idea of how the study is carried out is: Two analysis frameworks need to be set up. One for PAD and one for ME. PAD is measured in available KPIs. In order to find KPIs to benchmark with, a branch standard document specifying processes and recommended KPIs in a nuclear generation plant, is found. From this document a quantity of technical KPIs is elicited and serves as a reference model. This model shows possible and desirable KPIs from which a nuclear power plant should be controlled. The more of the elicited KPIs a system can provide the higher level of PAD the system is said to have.

A framework for system ME is also developed. With the framework a system can be measured in terms of needed effort to perform a certain change on a certain set of systems. Different possible future scenarios with the possibilities to reach different levels of PAD and ME are developed. For each scenario the PAD and PAD is measured at the unit of study. That is; every scenario represents a possible future that may be more or less hard to get to and may generate more or less of the elicited KPIs.

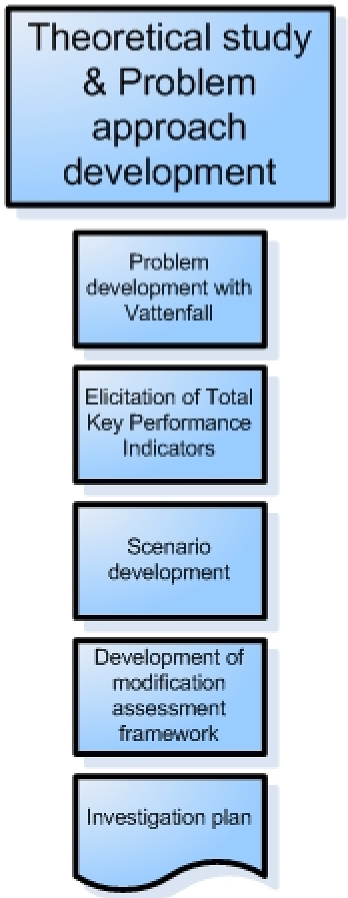


Figure 3, Theoretical study & Problem approach development and its elements.

3.1.3 Empirics

As introduced in the previous section the Empirics has mainly two different aims: Measuring PAD and measuring ME, for future scenarios. This is carried out first by letting the Business Information Officer at the utility company prioritize the processes and sub-processes associated to the elicited KPIs in order to provide a plant specific weighting of the nuclear process standard making different KPIs different important in the evaluation of different scenarios. Learning about which KPIs the power plant provide today made it possible to assess the PAD of today. For each of the future scenarios, the PAD can be assessed by the number of the elicited KPIs, that the scenario would provide. Then finding out, through which systems the KPIs are produced, the Modifiability of these systems is measured. From the set of future scenarios a Change complexity of each scenario can be assessed. The Modifiability and the Change complexity construct the ME for each scenario.

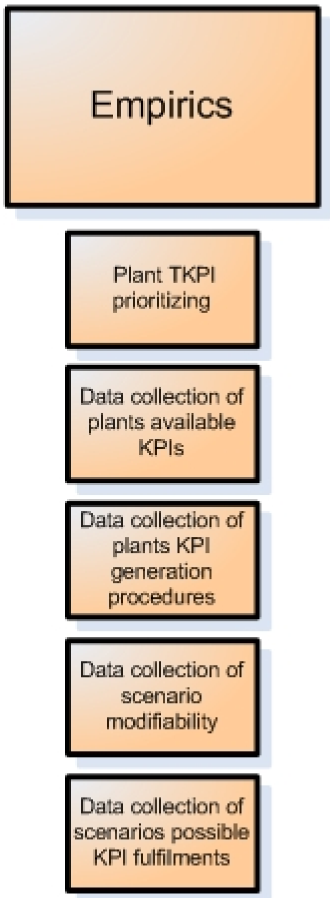


Figure 4, Empirics and its elements.

3.1.4 Analysis

Following the Empirics the Analysis is carried out. For each future scenario a comparison is performed between the KPIs generated by the scenario and the SNPM KPIs with respect to priorities made by the power plant. For each future scenario a measurement is carried out of the modification effort needed to reach the scenario's system solution from the system solution of today. Finally these two are brought together giving a result showing the best scenario choice that could be made maximising the Business Intelligence value and minimising the modification efforts.

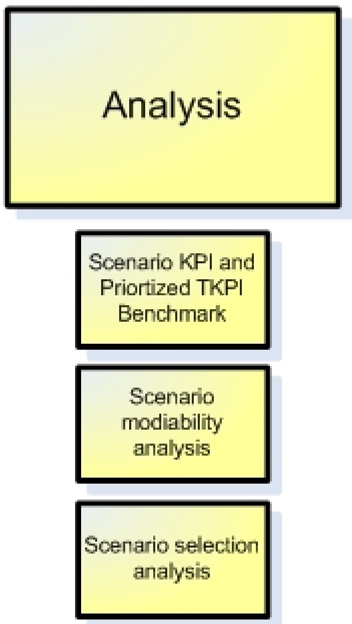


Figure 5, Analysis and its elements.

3.1.5 Presentation and Project closing

The last phase of the project consists of communicating the results to the interested parties in the best possible way. A Final Report including everything and a much shorter Executive summary communicating only the most important information and the studies result is written.

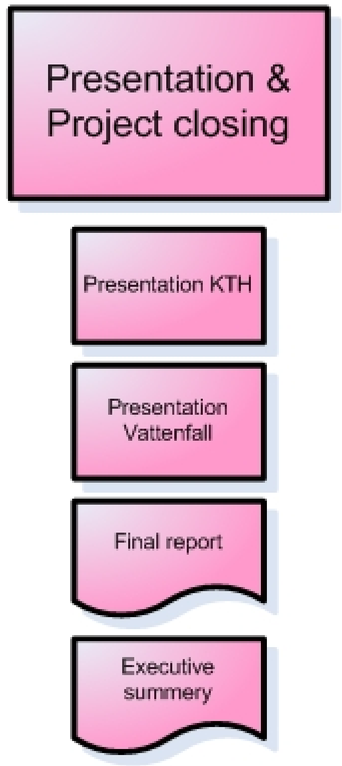


Figure 6, Presentation & Project closing and its elements.

3.2 Strategy

Answers to the questions of the study is provided through a correct analysis that is based on valid and reliable data [54]. The key word in these sentences is data. The empiric data is depending on outside parties i.e. in order to achieve good, valid and reliable results, data from external parties needs to be collected. In a huge organisation as the utility company, this is much easier said then done. Just the geographical distances between many of the different business units have made the planning of interviews harder.

During the entire process a lot of work and focus has been put on the arrangement of contacts. The aim was always to take care of the contacts with the utility company as soon as possible and the contact work enjoyed a special priority towards other tasks.

3.3 Striving for quantitative study

There are two approaches of a study, either it can be quantitative or qualitative. The quantitative method is used when information can be represented in quantities, i.e. they are measurable. The work initially starts with setting up strategies, identifying success factors, stock holder wishes and critical functions, then suitable quantities is determined. The study normally is performed with enquiries or interviews, a great effort should be put to the formulation of the questions. [49]

A qualitative method is more suitable when underlying needs, expectations, wishes and apprehensions is to be studied. Normally this is done through deep interviews, focus groups, observation and projective methods. [49]

This study strives towards a quantitative approach and the theoretical framework and the analysis method are built up in a strictly quantitative manner. The reader will find that in the chapter *4 Theory* and *6 Analysis* that the three frameworks being the foundation for the whole study is a built up in a way that quantified measures are summed together forming a quantified result.

There is though, however, some qualitative parts of the study. To a large extent the theoretical research performed to build these quantitative frameworks are done in a qualitative way where huge amounts of non-quantified information has been interpreted in a qualitative way and well-founded conclusions have been made. A strictly quantitative study would have to build its result on quantitative sources of evidence, preferably enquiries. While this study sometimes interprets qualitative answers from the interviews into quantitative that fit into the quantitative analysis frameworks. In fact, even some specific parts of the analysis framework are of more qualitative character. The best example is probably the elicitation of scenarios which is done, as part of the analysis phase, through qualitative interviews. See further in *3.5 Method for the Empirics* and *3.6 Method for the Analysis*.

3.4 Constructing theoretical framework

In this study two types of theory are required, first theory used for assessing the PAD in the systems and secondly there is the matter of assessing the ME which consists of Modifiability and Change complexity.

The theory is modelled as Extended Influence Diagrams (EID). An Extended Influence Diagram is, briefly described, a directed graph which has the ability to model uncertain variables and decisions [24]. The EID allows for a lucid break down of the studied topic in its integral parts so that the context becomes clearer. It also helps in identifying factors applicable to the study that are unambiguously measurable.

In the following passages it will be described how information has been extracted in order to create the theoretical frameworks. For the interested reader more extensive information about how the EID work process has been developed and used can be found in *Appendix C*. For the very interested reader even more information about EIDs and how they can be generated can be found in [24, 31].

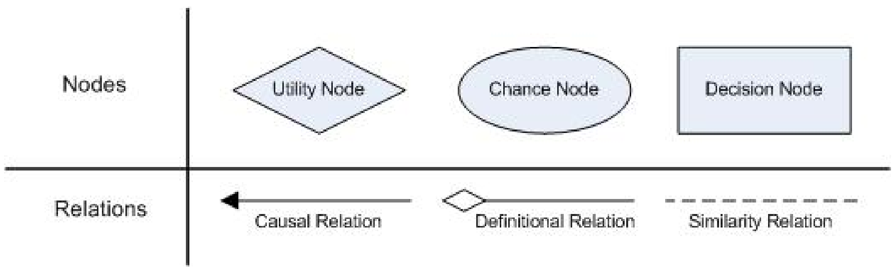


Figure 7, example of the notation of EIDs.

3.4.1 Building the ME framework

In order to assess certain ME when changing the systems of today into one of the future scenarios we need an EID framework. Due to the many relations between Modifiability and Maintainability, such a framework can be produced from a maintainability framework.

This study has benefited from three other maintainability frameworks: [23, 29, 42]. The creating of the framework has been carried out by the following persons: Robert Lagerström PhD student at the Royal Institute of Technology (KTH) and Göran Lindmark, Johan Ullberg and Jonas Öhrström, all master thesis students from KTH writing about modifiability on the utility company. By building a Maintainability framework together with others the theoretical framework of this study has been allowed to become both broader and deeper than the timeframe for a single master thesis study could allow.

Constructing the Modification effort EID framework

The processes of building a Modification effort framework can be divided into three different phases:

- Research of new articles and construction of EIDs
- Maintainability EID´s consolidation
- Refinement process when selecting change related parts from consolidated EID, ISO Standard Document, an early summarising article and some additional articles

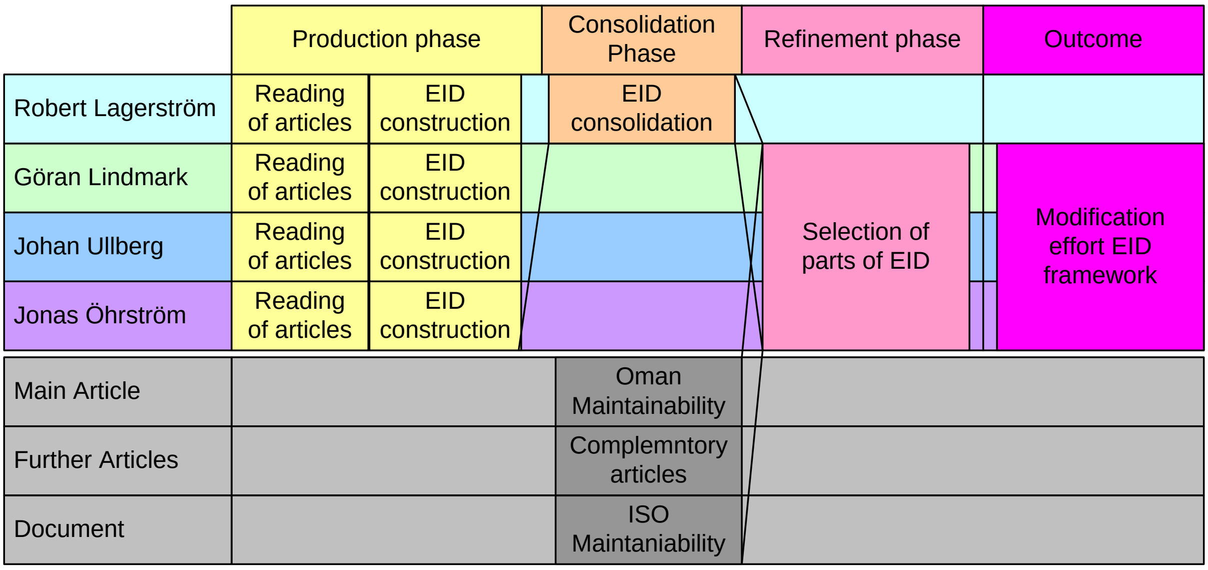


Figure 8, Process of building a Modification effort framework.

The first phase is performed together by all. Each one has read many articles and has then selected between 3 and 6 articles most fitted to create EIDs from. The process of creating EIDs is described in [31] and even more detailed in *Appendix C*.

The second phase has been done by Robert Lagerström based on the Maintainability EID´s from the first phase and in accordance with the consolidation processes described in [30].

The third phase has been done together by three of the participants and is based on consolidated EID, an ISO Standard Document [23], Omans article about maintainability that summarises the work done on maintainability until 1991 [42] and some other complementary work [2, 19, 20, 36, 37]. The outcome of the refinement process is the Modification effort EID framework which one can find in *6.4 Modifiability*.

Framework weights

There are several different ways of given weights to the theoretical frameworks and a brief description of some that has been considered for the modification effort framework is described here.

Intuitive weights

One approach would be to let the researcher intuitively decide which aspects are more important and which are less important. This approach is easy to carry out but inflicts high demands on the researchers regarding their competence within the area and ability to make these decisions unbiased.

Expert weights

If the required competence can't be found among the researchers another possibility would be to let a well-reputed expert set the weights. This is often a good strategy if it is possible to find such and expert willing to perform the work. The disadvantage is as in the case of intuitive weighting that it will be the subjective opinion of the expert that is modeled. A possible way of working around this is to involve several experts but it is of course even harder to find several experts than one expert.

Commissioner weights

If the study is commissioned by a person or organisation and they have some insights in the area it is possible to ask the commissioner to estimate how the nodes should be ranked. This has the advantage that the commissioner, hopefully, knows which areas that is important for the organisation and the study will therefore be performed more in accordance to this. Disadvantages are essentially the same as for the two previous methods.

Reference weights

Since the theory in this study will be built from literature it would be possible to assign weights with the help of the read articles. One approach is to count the number of references that mentioned each of the attributes. An other is to also include how important each of these references are, for instance by looking at the number of citations each reference has. Using this approach could seem very objective but the researcher could have searched for many references to an attribute just in order to understand its meaning and such errors would lower the objectivity.

This study takes the approach of letting the commissioner rank the different attributes used both in the case for PAD part of the framework and for the ME part of the framework. The upsides mentioned earlier seemed compelling and the downside of being subjective was not considered such a big problem (in some ways this subjectivity is the strengths of the method since it focuses on what the commissioner wants). The other possible methods were rejected mainly because an expert was hard to locate, the researcher was lacking the depth of knowledge needed and the reference approach would suffer too much of the selective searching mentioned above.

The weights for the ME framework will therefore be collected together with the empirics itself, after each question the respondent will be asked a question where he will asses this questions relevance for the studied topic.

3.4.2 Building the PAD framework

Building an assessment model for Business Intelligence is extremely complex since the amount of possible properties are just as huge for a BI system as it is for any system. That to measure the BI of a system completely one would have to measure all aspects like, for example, the usability and security. Therefore a simplification and specialisation of the model according to the preferences of stakeholders is needed. [27]

In accordance to this, the utility company specified that the degree of process assessment and the availability of Key Performance Indicators are what they consider the most important parts of Business Intelligence and this study therefore restricts itself to study BI as a function of Processes Assessment Degree (PAD). The first thing is therefore to create a reference model of KPIs serving as a framework for measuring BI.

3.4.2.1 Desirable Technical Energy Generation Key Performance Indicators Elicitation

In *3.1.2* above, a number of Key Performance Indicators that are so generally accepted that they have become branch standard is needed in order to assess the level of Business Intelligence at the utility company. As written about in *4.3 Nuclear Organisations* there are many organisations and committees working with these issues. To know what the branch consider standards, branch companies were asked.

Contributing companies

The companies that know best about Business Intelligence standards in the nuclear industry is, maybe surprisingly, not the nuclear energy companies, but rather companies selling software, that provides BI systems to nuclear energy companies. There are many such companies, both companies that are selling BI systems to all kinds of companies and companies with a specific orientation towards the power generation industry. In this study the following companies were searched for information about possible branch standard KPIs:

- SAP
- Oracle
- Indus
- MRO
- SPL
- RTPM

From the information found on their homepages and through contact by e-mailing it was possible to determine standards really recognised by the nuclear industry. [47]

Contributing documents

From the information given by the above mentioned companies, documents from two nuclear organisations were selected. The two documents mainly contributing to this study is WANOs 2005 Performance Indicators [51] and NEIs Standard Nuclear Performance Model – Revision 4 [39].

WANO is considered the top organisation, which connects different nuclear industries of the world. Their nine performance indicators have been assessed since long by a large part of the nuclear industry. This is very fundamental when it comes to performance indicators for nuclear industry. Or as Joseph Schippert, Lead Product Manager for the Nuclear Industry Solutions at MRO puts it: "*I can say from first-hand experience that these performance indicators are taken extremely seriously by everyone in the industry!*" [47].

The Standard Nuclear Performance Model was originally developed and issued in 1998 by NEI, EUCG and INPO with the aim to improve benchmarking effectiveness. It defines processes for nuclear power plants and connects a number of KPIs to these processes, thereby making it possible to measure the performance in each process. Furthermore it defines sub-processes and KPIs connecting to the sub-processes. Not only the three just mentioned organisations but 13 other organisations including EPRI are contributing to the standard and some carry special responsibility for certain sub-processes. The extensive document describes in detail how each KPI can be measured, but is unfortunately not complete. The SNPM is continuously being developed as different contributors agreed upon what KPIs should be used to measure a certain process or sub-process. Both more processes and KPIs are presented as new revisions of the document are issued. [39]

The great advantage of these two documents is that they are contributed to by practically every well-renowned organisation in the nuclear industry. It would be hard to claim that the list of KPIs founded in either of the two documents is anything but industry standard since all "the big ones" stand behind it. On the other hand it is gravely shifting in completeness. While some parts are well covered other processes totally lack KPI measures. The SNPM document alone is the basis for the PAD framework and the WANO list is used for the understanding of certain KPIs in SNPM.

3.4.2.2 The two level framework

In the SNPM both processes and sub-processes are recognised but the processes are not a function of its sub-processes. In the same way the KPIs of the sub-processes do not sum up to the KPIs of their related process. And even if, in most cases, a correlation between the process KPIs and its sub-processes KPIs there is in general no direct connection.

What this means in practice is that building the PAD framework as a tree where processes has sub-processes as "parents" with KPIs as leaves will prove to be inaccurate with the SNPM. This opens for a separate measurement, where sub-processes and processes are measured separately and both contribute to the Business Intelligence business value directly as can be seen in *4.6 PAD framework*.

3.4.2.3 Key Performance Indicators preference

All of the processes and sub-processes are important in nuclear generation [39]. But some of them might be more important to control and measure in a PAD perspective. The importance is the weights of the BI framework and just as in the modification framework the process assessment degree framework is weighted with a commissioner weight method. This is because what is important to assess may depend on how the process is carried out locally, since different ways of carrying out a process may influence how hard it is to control and how much measuring is needed. Therefore it is suitable that the prioritization of the processes and sub-processes is done locally at the unit of interest.

Performer of KPI Prioritization

For the purpose of doing the prioritization of both processes and sub-processes, the Business Information Officer at the utility company, was contacted. He agreed to do the prioritization and contributed thereby with the weights to the PAD EID framework, see further *4.6 PAD framework*.

Achieving the prioritisation by Focal Point

By letting the Business Information Officer do pair wise comparisons between the processes and sub-processes, the prioritization was achieved. For this, a web-based tool, called Focal Point, was used. This tool allows prioritisation between large quantities of objects. In this case the objects were processes and sub-processes. While prioritising, a question of comparison is shown together with two of the objects. It is then possible to make a comparison which object fulfils the given question the best ranked on a nine level scale, see *Figure 9*. When the given comparison between two objects has been answered a new comparison between new objects shows up. By comparing the objects in pairs on the nine unit scale it's possible to prioritise all objects with respect to each other.

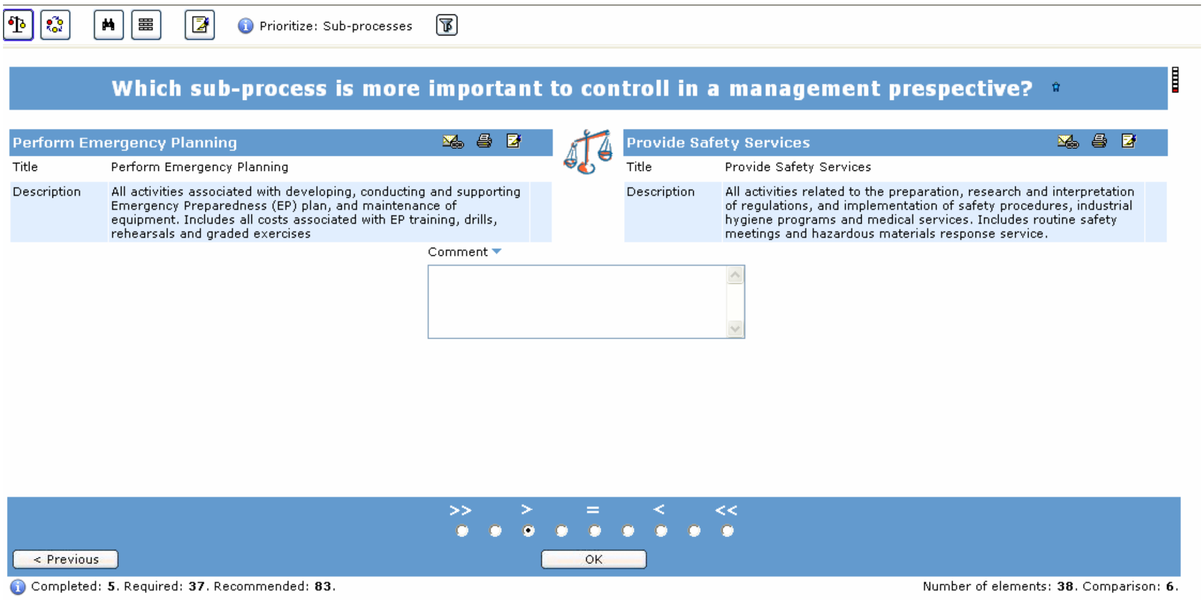


Figure 9, Example of a Focal Point comparison.

The pair wise comparisons were done by using the following two questions;

For the processes;

Which process is more important to control from a management perspective?

And for the sub-processes;

Which sub-process is more important to control from a management perspective?

Focal Points handle the mathematics and keep the number of necessary comparisons to a minimum. For a good result with a high certainty of the consistency, a number of 2*n+1 comparisons are needed. Where n is the number of objects that should be compared.

3.5 Method for the Empirics

As mentioned in *3.2 Strategy* the good data collection during the Empiric phase was the real key to the success of the study. In this chapter a more detailed description of the methods used for achieving this data is given.

3.5.1 Sources of evidence

Yin mentions six different sources of evidence as being the most important data collection methods: [54]

1. **Documentation**
   Documentation is a very common source of information and is probably used in almost all case studies. Examples of documentation is: Letters, notes, calendars, meeting notifications, action reports, questionings, progress reports and other studies or estimates as well as press cuttings or extracts from newsletters. [54]

2. **Archived records**
   Archived records in the form of, for example, data register or files could also be relevant for some studies. Examples of archived records are: Maps, Lists, Examination data, personal registers, service registers and organisation registers. [54]

3. **Interview**
   One of the most important sources of evidence for case studies are interviews. The interviews should be more of guided conversations than interrogations. The most probable is that even if a consistent line of questions exist, the case study interview will give answer to a lot more than the asked questions. [54] See further *3.5.3 Method for interviews* below.

4. **Direct observations**
   By making a visit on the unit of interest for the study an opportunity to directly observe interesting behaviours, activities and processes is given [54]. Direct observations also makes it possible for the researcher to get direct input that has not been "filtered" by anyone else which be the case for interviews.

5. **Participating observations**
   The participating observation is a special case of the direct observation. The difference is that the researcher takes a role with in the range of what is to be studied [54]. For example could the role as a project member in development project be adopted to give a unique insight in how certain types of projects are carried out.

6. **Physical artefacts**
   The last source of evidence that normally could be used in a case study is a physical or cultural artefact, technical device or an instrument or some other physical object. Physical artefacts normally play a not so important role for most case studies. [54]

3.5.2 Choice of sources

In order to select the best sources of evidence, a Value/Cost estimation was performed. A Value/Cost estimation sees to which value a certain source of evidence would provide the study and which cost it would entail, measured in work effort. [16]

Sources of evidence Value/Cost analysis	Value	Cost	V/C
Documentation	6	4	1,5
Archived records	6	8	0,75
Interviews	8	4	2
Direct observations	8	8	1
Participating observations	3	10	0,3
Physical artefacts	1	8	0,13

Table 2, Sources of evidence Value/Cost analysis, scale from 1 to 10.

The result is the quota between the value and the cost and gives a hint about which sources that might be useful. Following this analysis Interviews was selected as the main source of evidence and Documentation as a supportive source. As Interviews was the main source of evidence, a more detailed look into the interview methods will follow.

3.5.3 Method for interviews

Since the interviews was the most important source of evidence a little more details about how they were carried out has been developed.

Interviews can be differing in many ways and the basic issues that have to be decided about are: The Form, the Constellation and the Type of Questions [34]. Furthermore there are many things which are important to think of when conducting an interview in order to make sure to get what is wanted and interview policy have been put together.

Form

In this study, face-to-face interviews have been preferred in preference to an interview conducted over telephone or by email. So that as much information as possible could be drawn from the very few interviews that was granted and to make it possible to ask follow up questions. Face-to-face interviews was also by far the best way of understanding subtle messages that were not expressed literally [34]. If a face-to-face interview could not have been given because of scheduling issues a telephone interview would also have been considered as an acceptable form for this study.

Constellation

Interviews can be conducted in group or individually and it is possible to interview a whole group or one person at a time. In this study interviews with both one person at a time and a group at the same time was used and (naturally) with only one interviewer. The advantage with interviewing only one person at a time is that it is possible to go deeper and personal opinions will be raised [34]. The disadvantage with being only one interviewer is that it's harder to both guide the conversation and take record of the answer.

Type of questions

Questions can either be of a structured type when more quantitative answers are wanted, or less structured when the questions are more open and the interviewee are allowed to give more elaborate answers. Interviews can also be unstructured with only qualitative questions and with more anecdotic answers [34, 49]. This study have gone the middle way allowing itself to ask qualitative questions to get general ideas how things work or how different parts connect and quantitative questions that provided the data for the theoretical frameworks.

Policy for a successful interview

Since the number of sources of evidence in this study was very limited a lot of thought was done over what issues are critical for the success on an interview. The following list form a policy that was used during the entire empirical process and was put together from the advice given by Merriam and Yin. [34, 54]

- Be well prepared.
- Prepare the one that will be interviewed
- Listen carefully and be objective to the information
- Be flexible and open for new angle of approaches that comes up during the interview
- Identify and interview the key information holders
- Identify contradictions immediately and ask follow up questions

3.5.4 Principles for the Empirics

According to Yin there are three important principles that should be applied for all collection of data. By following these the study enjoys an increased validity and reliability: [54]

1. Use different sources of evidence
2. Create a case study database
3. Maintain the chain of evidence

Studies that uses many sources of evidence are ranked, in terms of overall quality, higher then studies with just one kind of information [54]. In this study however, the data collection was dominated by interviews and therefore suffers from the risk of lacking validity. Throughout the empiric process, it was therefore important that a strong focus lie on finding system documentation and other documents, as mentioned in *3.5.2 Choice of sources,* in order to back up the information given during the interviews.

A well-kept study database is the second principle and in order to achieve it the following rules was set up and followed for the empiric data:

- Create a reference system to all produced material as well as empiric-collected data that allows a simple distinguishing between different documents.
- Save all material in electronic form, the data should be continuously, on weekly basis, be made a back up of on a separately located hard drive.
- Keep a well updated reference list over all important literature that has been found useful for the study.

Finally, the reliability of the study was obtained by maintaining the chain of evidence. This was done by clearly showing from where information comes and by documenting the working process so that it became possible to track back a result through the analysis back to the source of evidence. Well-written reports with good references and weekly reports about the work progress ensured the reliability of the case study.

3.5.5 Study objects

According two Yin the basic design of the case study can be of four different kinds: [54]

1. The single-case holistic design
2. The single-case embedded design
3. The multiple-case holistic design
4. The multiple-case embedded design

The single-case study is normally not preferable since it doesn't enjoy the same robustness nor give a basis for statistical analyses as a multiple-case study. There is though, however, according to Yin, many reasons that could justify a single-case study. A holistic design means that there is one studied unit only while an embedded approach looks in to many units of analysis. An embedded approach is good when subunits is easy to identify and makes benchmarking between the units possible. [54]

3.5.6 Choice of study object

This study was performed in a of single-case holistic design. This means that the study was done with a single case and only one unit of study was identified. This was mainly because of the time restrictions.

After dialogs with both the mentor and sponsor at the utility company but also with responsible persons at two nuclear power plants about which study object would be suitable for this study and where time for the study could be placed for disposal, a nuclear power generation plant was chosen. The Business Information Officer at the utility company was the main supporter of this decision.

The study of Business Intelligence systems at the nuclear power plant can be seen as a pilot study and many similar studies could easily be carried out elsewhere. Especially at another nuclear power plant where the same frameworks could be applied without any great changes.

3.5.7 Empirics at the nuclear power plant

A workgroup of people from different areas was set up at the nuclear power plant in order to collect empirics. This group included decision makers in both IT and Business fields. The workgroup had experiences from different areas, so that most of the KPIs given in the SNPM and their supportive systems could be identified, if existing within the organisation of the nuclear power plant. Furthermore documentation about the systems in question was used to back up the information given in the interviews.

Secondly a number of interviews were held with IT personal or consultants in order to determine the modifiability in the systems. Those who were in question were making modifications and maintenance work on the systems.

See further in *5.1 Interviews* for the performed interviews.

3.6 Method for the Analysis

For the purpose of analysing the collected data the theoretical frameworks will be used. The framework is constructed in the form of an EID as described in *3.4 Constructing theoretical framework*. The EID framework comes with a well developed method for analysis [24].

The EID framework is based on Bayesian networks, where probabilities higher up in the hierarchy are based on the probabilities of its sub-nodes. The higher node is called "child" and the sub-nodes are called "parents" to the child node. This builds up the framework, where each child node has a set of one or more parent nodes until the end-nodes, which have no parents. There are, as can be seen in the framework appendixes; *Appendix A and Appendix B*, very common with multiple levels of nodes, which means that a node could be both child to one node and parent to another at the same time. [24]

Each node could be in one out of a set of certain states and the chance that a node is in a certain state is shown as a state probability. For the end nodes, these probabilities come from answers given in interviews or from other kinds of sources of evidence. For the child nodes the probabilities depend on the state probabilities of their parents. Each parent could affect its child differently and also differently much. That means that some parent nodes could be more important for the state of the child node(s) then others. This is called weight.

This chapter will describe how the general ideas about how end-node probabilities has been set, how the uncertainty of a node's state is handled, how child node state probabilities is determined and how weights between parent nodes work. Three different sub-chapters of *6 Analysis* are then describing more specifically how special analysis method adjustments have been applied on the three main parts of the theoretical framework because of the unique circumstances of each one of them.

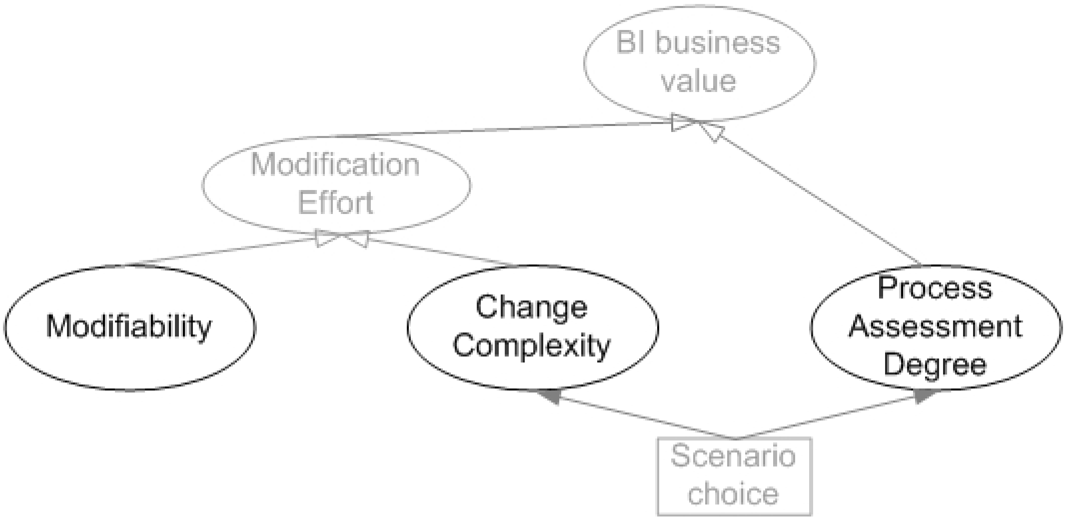


Figure 10, The main structure of the theoretical framework, the three different main parts are darker.

3.6.1 Probabilities of end nodes and credibility of sources of evidence

The end nodes of the Extended Influence Diagrams, or the frameworks as they are normally called in this thesis, are assessed through empiric research. The EID analysis method requires that all collected data are quantitative. Different nodes may be quantified in different states. The only requirement is that they are quantified into a finite number of states for example 5 states (*Bad*, *Poor*, *Ok*, *Good* and *Excellent*). A node could though have a set of probabilities of obtaining many of the different states. Given the example of possible states a node may for example have 25 % chance to be *Ok*, 75 % chance to be *Good*.

For the simplicity of the analysis, all nodes in this study has been quantified into the same three states (*High*, *Medium* and *Low*) plus an extra state *Uncertain*, which handles the credibility of each answer, i.e. the credibility of each end-node state. A node in the framework may then be given its state probability in an interview, which is considered to have a credibility of 80 %. If the answer corresponds to the state *High* it would have the probabilities 80 % *High* and 20 % *Uncertain*. The number of possible states is kept low, since the underlying mathematics for the analysis is drastically increased by an increasing number of states[24].

For each end node in the EID frameworks a question will be asked during an interview, that gives enough data to assign one of the three possible states to the node. It is also possible that a node's credibility might be strengthened by other kinds of sources of evidence. See further in *3.5.1 Sources of evidence*.

In case two sources point towards different states for the same node the probabilities split up on the different states. So, for example, if another interview is done from the previous example with another 80 % credibility answer, which corresponded to the *Medium* state, the node would be given 40 % probability of *High*, 40 % probability of *Medium* and 20 % probability of *Uncertain*. For more specific explanation how the credibility and the probabilities has been handled in each case, please see the sections *6.1*, *6.4* and *6.5* below describing the specific analysis applied for each of the different framework pieces.

In case many sources give data to determine the state of a node, credibility could either be calculated as an average of all sources or be considered higher if all sources point to the same state and then lower if they point to different states. In this study the first of the two ways of handling many sources has been used. Sources in this study seldom look upon a node with the same underlying system in mind when answering the node's question and quite often the different systems have proved to be working differently and therefore giving different answers to the same question. It would be wrong to say that these questions' answers have a lower credibility.

Generally the credibility is calculated from three elements. The first two are credibility estimates done during interviews and the third is a credibility of the source of evidence in the cases interviews have been complemented by a visual proof. In the most common case, only interviews has been used which is considered having a base credibility of 90 %. Interviews could though, have a lower credibility, depending on the average of the two credibility estimates. The first credibility estimate was done by the interviewer and shows how sure the interviewee seems to be on his/her answer according to the interviewer. The second is done by the interviewee himself/herself also ranking how reliable the answer he/she has just given is. The ranks in both cases are between 1 – 3, where a 1 is considered almost guessing and a 3 is considered completely sure. The table gives an overview of the values for credibility normally used:

Credibility	Interview avg. cred. est.	3	2,5	2	1,5	1	
Interview		90%	80%	70%	60%	50%	
Visual proof							100%

Table 3, credibility depends on both the kind of evidence and in the case of interviews on the average of the credibility estimations.

A different approach to handle credibility, which also has been tried out is to exclude the *Uncertain* state and distribute the uncertainty on the other states. If the Credibility is 70 %, 30 % would then be distributed over the other three states giving 10% to each. In this case the Credibility of the whole study would be calculated as the standard deviation of the top-node. This approach has proven to give a preposterous huge uncertainty for just very small uncertainties in the end-nodes. This is because a child with two parents where one is 100 % *High* and the other is 100 % *Low* does not become 100% *Medium* rather 50 % *High* and 50 % *Low*. This would cause a great uncertainty in the child node. With the current way of handling Credibility, the Uncertainty of the top-node is directly dependent on the credibility of the end nodes and how much they weight compared to other end nodes. The uncertainty of the whole study can be directly read from the probability of the *Uncertain* state of the top-node.

3.6.2 Weights and probabilities of non-end nodes

In order to set the state of a child node it is necessary to know the following:

- Which are the parent nodes?
- What probabilities of states do the parents have?
- How does a certain state of parent affect the child?
- How important is each one of the parents for the state of the child i.e. which internal weight lies between the parents of a child node?

The first question comes with the theoretical framework. It clearly shows which parent nodes each child node has. Please note though that EIDs are not necessarily tree structures since parents can have more than one child.

The state probabilities of the parents are known in the case of end-node parents as described above and in case the parent is not an end-node it should have been previously calculated from its own parents. Note that this means that the framework is always calculated from the end-nodes and upwards.

The state of the parents may affect a child differently. In this study however the understanding of framework is greatly simplified by the fact that all nodes have the same states. Generally the state *High* of a parent adds to the *High* state of its child. It works the same with the *Uncertain* state. There are exceptions however between the different parts of the frameworks. A *High* Modifiability leads to a *Low* Modification Effort.

The weights are the last important piece when determining the state probabilities of a child node. The weights for each end node has been elicited a bit differently for each of the three framework parts and will be further discussed in the following framework-part-specific method analysis chapters. Generally each parent node carries a weight which can be compared with the other parent's weights. The normalised value of the parent weights shows how much each parent will contribute to the child. This is best illustrated by the following example:

Parent 1 has the State *Low* with 80 % probability and *Uncertain* with a 20 % probability and a weight of 12.

Parent 2 has the State *High* with 80 % probability and *Uncertain* with a 40 % probability and a weight of 4.

In this case the normalised weights between the two parents would be Parent 1: 0.75 and Parent 2: 0.25. The probability tables for the two parents would be

Parent 1	
High	0
Medium	0
Low	0.8
Uncertain	0.2

Parent 2	
High	0.6
Medium	0
Low	0
Uncertain	0.4

Table 4 and Table 5, probability tables of Parent 1 & 2.

The example framework now looks like this:

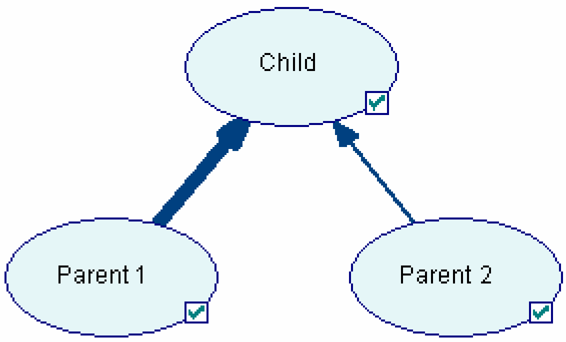


Figure 11, example framework child and parent structure. The greater arrow from Parent 1 indicates it's higher weight compared to Parent 2.

In this example a state *High* in a parent adds probability to the state *High* in the child. The same counts for the other states. This means that the following dependency table which shows how certain states of the parents give certain probabilities to Child directly reflects the weights of each parent:

Parent 1	High				Medium				Low				Uncertain			
Parent 2	H	M	L	U	H	M	L	U	H	M	L	U	H	M	L	U
High	1	0.75	0.75	0.75	0.25	0	0	0	0.25	0	0	0	0.25	0	0	0
Medium	0	0.25	0	0	0.75	1	0.75	0.75	0	0.25	0	0	0	0.25	0	0
Low	0	0	0.25	0	0	0	0.25	0	0.75	0.75	1	0.75	0	0	0.25	0
Uncertain	0	0	0	0.25	0	0	0	0.25	0	0	0	0.25	0.75	0.75	0.75	1

Table 6, dependency table of Child reflects the weights of Parent 1 & 2.

The given probabilities of each parent and the normalised weight are simply multiplied in order to find each parent's contribution to the probabilities of the child.

Parent 1 contributes 0.8 * 0.75 = 0.6 to the *Low* state of Child and 0.2 * 0.75 = 0.15 to the *Uncertain* state of Child.

Parent 2 contributes 0.6 * 0.25 = 0.15 to the *High* state of Child and 0.4 * 0.25 = 0.10 to the *Uncertain* state of Child.

This leads to the resulting probability table of Child:

Child	
High	0.15
Medium	0
Low	0.6
Uncertain	0.25

Table 7, the probability table of Child.

This example will finish the two general chapters on *Method for the Analysis* and more specific details on how empiric data is treated for each framework part will be given in *6 Analysis*. It is also important to understand the idea about each part's place in the framework and how they connect with each other and therefore a short chapter on framework-part connections is also given there.

4 Theory

In the beginning of this chapter a background for the Modifiability and BI topics are given so that the reader unfamiliar with these subjects gets a clear picture what it is about. This background information is the base for the future systems scenarios that also are presented in this chapter.

4.1 Modifiability

It's all about how easy it is to modify a certain system. But when looking closer at it, the factors that could influence the modifiability are endless. For example the complexity of code might not be as an important factor as on what platform a system has been built or the experiences of those performing the modification in some cases. In some cases organisational opposition can be the greatest factor for having problems implementing a modification, while in other cases the bad project and work model are what causes problem.

Yet another issue is the many synonyms that can be found in literature. Flexibility and maintainability are often referred to as something very similar to modifiability. A good example is this definition of maintenance: [53]

The process of modifying a software system or component after delivery, to correct faults, improve performance or other attributes, or adapt to a changed environment.

The need for a clear definition which states what should be included in the Modifiability concept is evident.

4.2 Business Intelligence

Business Intelligence (BI) is one of the most important concepts of this master thesis and also one of the technologies that have proved to give the best return of investment when it comes to competitive information management technology [9]. While BI springs from very old ideas the technological implementation and practical usage is fairly new and still under rapid development. [18]

For this study, as can be seen further in *4.6 PAD framework* only a small part of what really makes a whole Business Intelligence system will be quantitatively measured. Specifically this study has chosen to focus on the existence and usage of Key Performance Indicators (KPIs), a concept that soon will be well explained.

Even though this study focuses on KPIs, it is important to understand where KPIs fit into the BI system and also what the BI system aims to deliver and what parts are essential for it's success. Therefore a short introduction will first be given to BI itself and the main components that build up the future scenarios. Information on other interesting related technical topics can be found in *Appendix D*.

4.2.1 The concept of Business Intelligence

The concept of Business Intelligence has developed from the traditional business environment analysis and the general idea is to gather information from different parts from inside and outside the business to create a better decision basis in order to become more competitive [43]. The most common definition of the concept is: [18]

The process of turning data into information and information into knowledge.

Some have also tried to develop the concept even further by talking about how to turn knowledge into intelligence and intelligence into wisdom and even wisdom into truth [50]. In order to understand the concept even better the following definition of information: [4]

Information = Data + Definition + Presentation

Already in 1958 H.P.Luhn wrote in the IBM Journal about what he called a Business Intelligence System that should "*utilise data-processing machines for auto-abstracting and auto-encoding of documents and for creating interest profiles for each of the action points in the organisation"* [32]. And this is still, almost 50 years later, more or less, what it is all about. The main development of Business Intelligence began in the beginning of the 90's and derived from managers' needs for effective analyses of enterprise data and by the soaring capacity of information technology and the negative price development [18, 43]. Soon a BI-technology became an essential element in achieving success in the ever more competitive environments of globalisation [28]. Today BI-systems exist in all larger companies and continue to develop swiftly. Many challenges still remain such as shortening the time from data to decision, implementation of artificial intelligence to handle the large quantities of information and BI-system modifiability in order to adapt better to business processes [18, 41, 43].

4.2.2 KPI – Key Performance Indicators

Key Performance Indicators are quantifiable measurements that reflect the critical success factors (CSF) of an organisation. They can either measure the performance of a process, a project or a whole business. [6, 39, 44]

The purpose of KPIs is to give a "to the point analysis" of the state of the process, project or business. KPIs should both reflect and help achieving the organisation's goals and also enable the organisation to benchmark itself with other organisations. [6, 44]. Correctly identified KPIs that are well aligned with the company strategies will provide a visibility into how well the company is maintaining it's strategic focus [5]. The concept of KPIs is relatively new and has fast developed to become one of the most important parts of performance assessment in the field of Business Intelligence.

In this study KPIs will be identified so that they reflect critical aspects of outputs of processes and sub-processes in nuclear power generation.

4.2.3 DW – Data Warehouse

In literature one can find slightly different views on what exactly should be included in the concept of the Data Warehouse (DW). One of the more cited definitions is W.H. Inmon's [22]:

A data warehouse is a subject-oriented, integrated, nonvolatile, and time-variant collection of data in support of management's decisions. The data warehouse contains granular corporate data.

Or maybe even a bit more tangible in R. Bose's definition [5]:

A DW is a structured extensible environment designed for the analysis of non-volatile data, logically and physically transformed from multiple source applications to align with business structure, updated and maintained for a long time period, expressed in simple business terms, and summarised for a quick analysis.

The most important attributes of a Data Warehouse can be derived from these two definitions.

- *Integrated*: A Data Warehouse is integrated, which signifies that data collected from different sources have to be converted to the same format before it is stored. According to Inmon this is the most important aspect of a Data Warehouse. [25]

- *Non-volatile*: In a Data Warehouse all data are unchangeable after having been integrated. While data in operative systems many times may be edited or erased, data in a Data Warehouse stay untouched. [25]
- *Time-variant*: In a Data Warehouse every piece of data is tagged with information about when it was stored in the database. This is also something which differentiates a DW from a more operative system, where the stored data always have to be up to date. [25]
- *Extensible*: A DW must have the ability to expand in order satisfy the demands to include either more data from the old applications or data from new applications.[5]

The main positive aspects of investing in a Data Warehouse are:

- *Competitive Knowledge Advantages*: Essential and earlier not known information can give the possessor of a well developed Data Warehouse enormous advantages. [22, 25, 50]
- *Higher efficiency among decision makers*: Data Warehouses integrate information from different systems of the organisation and create overall pictures of the organisation and provide bases for a decision. This dramatically shortens the time needed to do the analysis of the situation and to take the right decision. [25, 50]
- *Higher business opportunities perceptiveness*: A Data Warehouse helps company personal to identify hidden business opportunities through information dissimulation and data mining. [50]
- *Potentially high return on investment*: Even though the initial costs are high and the post installation maintainability costs are even higher, there are great chances that the investment will have a return in revenues to up to 600% as it will lower the cost of accessing information, improve customer responsiveness to company promotions and improve productivity. [22, 25, 50]

Identified downsides of making an investment in a Data Warehouse includes:

- *Very easy to underestimate data load time*: A very large part of the development time for a DW is taken by the time it takes to integrate data into the DW. The time necessary for this is very hard to estimate and very easy to underestimate. [25]
- *Hidden problems with the source systems*: The Data Warehouse might be very sensitive to errors in the source system, that provides the data. If there are errors generating wrong values or wrong relations between tables, these problems will probably be transferred to the DW. The problems might have worse consequences in the DW, than they had in the source system and it might take years, before they are discovered. [25]
- *Unclear ownership of data*: It might be hard to restrict the rights to access data and information, which should be kept secret. A DW brings great risks of spreading non-public information through the organisation. [25]
- *High maintenance*: A Data Warehouse requires huge amounts of maintenance. All organisational changes, which affect the company's processes, might imply large changes in the DW as well. [22, 25]
- *Long construction times*: The time needed in order to fully implement a DW is very long and can definitely be calculated in years rather than months. [25]

4.2.4 DM – Data Mart, dependent and independent

A Data Mart can be described as a small, department specific DW [55], but there are in fact a lot of attributes, that clearly separates the two concepts. First though, a definition:

A Data Mart is a subset of a data warehouse that supports the requirements of a particular department or a business function [26]

This to the point definition is today generally accepted, even though there are many issues of Data Marts, which still have not been settled. The main characteristics are:

- *Limited scope*: Data and the structure of data, that can be found in a DM is normally suited for the single department, which owns the DM [26]. Only the data which is needed for the own departments' reporting and processes controls will be found. Furthermore the structure of data is shaped by the end-user requirements into a form specifically suited for the needs of the department. [22]
- *Present data*: The amount of historic data saved in the DM also depends on the needs of the single department, which, to a great extent, means that almost no historical data is saved in a DM. All data in a DM should be current and present. [25]
- *Less complex*: The needs of complex data analysis, and applications pre-processing data is normally much less extended, than in a DW. Due to smaller amounts of data and the less complex structure normal database queries are normally used to extract information. [25] According to Inmon however it is the other way around; DMs are much more flexible and thus suitable for complex analysing tools compared to a DW [22].
- *Less costly*: The cost associate with saving and manipulating a single unit of data is much lower in a DM than a DW. [25]
- *Less restricted*: A DM, including its hardware, is normally owned by the department using it. This makes it much easier to modify and adjust, since the control systems and the rules restricting the use of a DW can be eluded. [25]

4.2.5 Combined DW and DM, Independent and Dependent

In most businesses DMs are much more frequently occurring then DWs. This is due to the relatively low initial cost of DMs, which make them the first try out. Later when a DW is built, the DM normally stays in place and works together with the DW. Normally two different cases of relation between the DM and DW is identified; the independent DM relation and the dependent DM relation. [17, 25, 50]

- The independent DM relation is characterised by the fact that the DM gets its information and operates completely independent of the DW. The DW might get its information from the DMs, but the DMs don't get information from the DW.
- The dependent DM relation is characterised by the fact that the DM gets all or at least some of its information from the DW. The Data Mart extracts the information it is interested in from the DW and processes it further.

4.2.6 What makes a good Business Intelligence system

The discussions on what makes a good Business Intelligence system have continued for as long as BI has been studied, which in practice means the last 20 years. Subjects have changed and as different theories have been tested proved working and not working. The concept of Business Intelligence has gone through an evolution from something very vague to a well developed technology with standards and best practices. Still, though, it is a technology with many question marks, and lots of theories untried.

The biggest BI question that arose in the nineties was concerning how and if the ideas about a company unique Data Warehouse should be implemented. Everybody understood the growing need for BI but few knew how to achieve it. For long time people believed in one of the two solutions; A BI system consisting of DMs or a BI system consisting of one DW, as described in 1997 by Chaudhuri and Dayal both working with R&D in Microsoft and Hewlett-Packard respectively [8]. Today this is hardly an issue anymore; soon it is common knowledge, that all larger organisations need an enterprise covering DW and that DWs and DMs should be seen as complimentary to each other. [18, 22]

In this chapter, however, no more classical questions will be discussed. Instead the best practices of today will be presented, showing the general ideas about the structure of a BI system in a large organisation. What is written here is what the research world of today finds the best way and what has actually been found working. [26]

The good BI system

The picture below describes the ideal BI system model in a larger company today, as it is described by many prominent experts. The composition of different parts will be explained and motivated since it is a very important issue to understand the basics, even though the technical perspectives of BI is not the main issue of this thesis. This model is mainly based on the ideas of [18, 22, 50] but support for different parts of it is founds in [1, 8, 17, 18, 22, 28, 41, 50, 55]. This model show the way data transforms to information and knowledge.

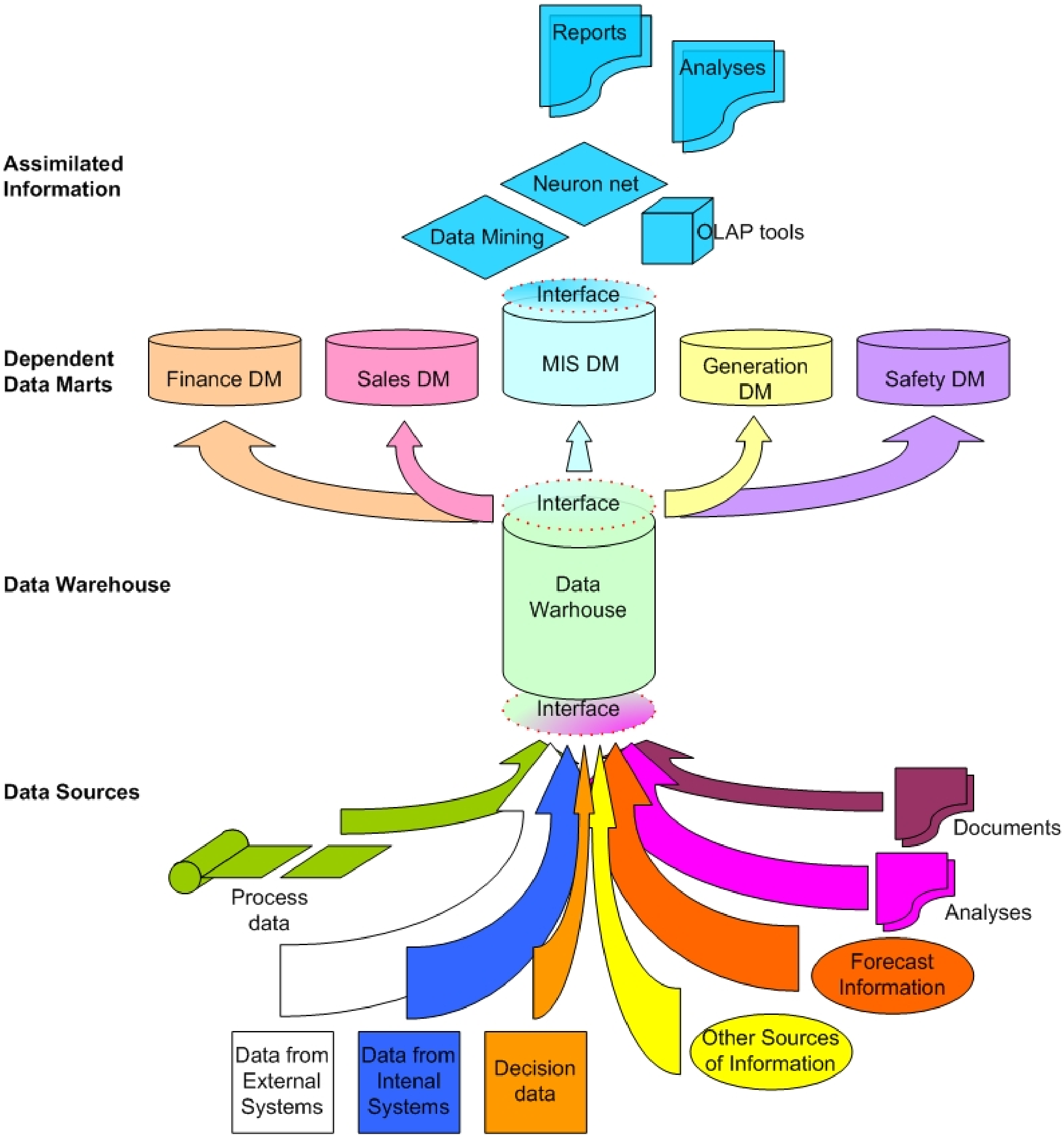
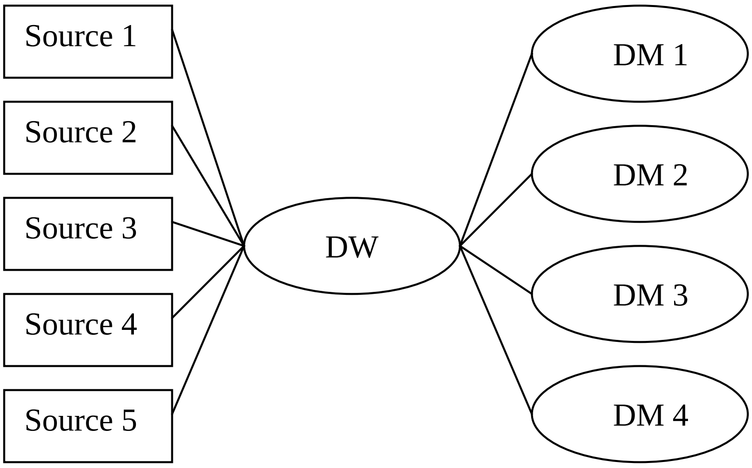
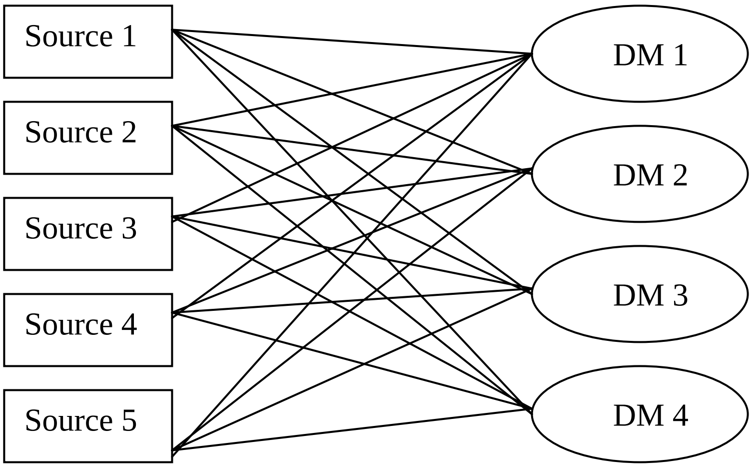


Figure 12, The concept of a good Business Intelligence system.

One of the most important concepts that can be found in this model is that of the dependent Data Marts. This means that the DMs collect all their data from the DW. This reduces the risks that different DM operates with different values of data. The DW becomes "the one truth". It also reduces the connection between different sources and the different DM to a minimum [22, 25]. This is perhaps best explained in the following picture.



Dependent DM



Independent DM

Figure 13, the different number of connections needed in a dependent DM vs. an independent DM.

In order to efficiently create a dependent Data Mart structure a great focus has to be put on the Data Warehouse. The DW is very different from a DM in its structure, even though many people traditionally have mistaken themselves to believe the two different concepts to be almost the same. But as Inmon writes: [22]

"Data warehouses are fundamentally different from data marts. The two do not mix – they are like oil and water."

This also stresses one of the most important issues of the Business Intelligence model; the interfaces. Not only is there a need of a well constructed DW, but in order to build BI it is also necessary to have clearly defined interfaces that makes it easy to hook on a wide variety of different kinds of sources to information. Process data from the process systems, data from other databases, data from strategic plans, as well as documents and others kinds of data need to have proper interfaces, so that the data can be stored correctly in the DW. Further more it is also very important that the DW provides interfaces on the other side towards the DM, so that the DM can extract data and information efficiently.

According to Inmon it is not a good idea to process data and use data mining techniques in the DW since the structure of the DW is normally not built for processing but rather just for storing and extracting [22]. Instead these techniques should be applied directly on the DM that has been constructed only to process data in the way that the department needs it. The DM can be more specifically structured towards certain extraction and analyse techniques than a DW could ever be.

4.2.7 The Key Performance Indicators in a BI system

Key Performance Indicators can be constructed from many different kinds of data. They may show how close an analysis came to the actual value in a certain process and they may consist of both process data and analysis data. Therefore the KPI very seldom can be seen as something coming from only one source of data.

Many different sources of data are stored in the DW in the model presented in the previous chapter. The information concerning a certain KPI is extracted to a DM.

In the MIS DM the data is stored in a way, which makes it easy to process. Data can now easily be processed and put together, so that it creates the KPI, which can be presented as information on a dashboard.

Data from totally different departments and areas of the organisation are all stored in a common DW. This makes it possible for every one to access all information about everybody and to use the information in any way to fit their needs. In this way information can be compared and verified and departments can easier benchmark against each other. Information will be treated equivalently from all sources and can therefore be easier understood by everyone. Of course it is always a question of security, who should be allowed to what information and some restrictions are naturally needed in most organisations.

4.2.8 Motivation for using Key Performance Indicators

The last ten years business management has changed from both in a technological and in an organisational respect [18]. Organisationally companies have become more process-oriented, trying to cut cost by adopting strategies to synchronise all business activities. These strategies need goals and metric-driven management in order to be enforced [18]. “*If you can’t measure, you can’t manage.*” [50]. With good Business Intelligence systems managers can learn how their business is developing at any time, instead of just learning about it at the quarter’s end [5].

If requires managers to ensure that the performances in all processes are effectively and continuously measured by KPIs [18]. Setting out and sharing goals and measurements at all company levels helps enforcing and communicating the strategy of the company to the whole organisation. That is, translating data into a detailed set of indicators, that are closer to the operational tasks, allows employees at all levels to better understand, what is wanted from them. [18]

Furthermore, Assessment Management and the use of standardised performance indicators make it possible to benchmark different companies and departments and to identify areas of possible improvement. The use of performance indicators has proven to be one of the most efficient ways of improving the overall quality the nuclear industry [39].

4.2.9 PAD measurement

As mentioned in *2.7 Assumptions* this study simplifies the complex concept of Business Intelligence a bit, when estimating the level of Business Intelligence in a nuclear power plant by measuring PAD; that is the number of available KPIs from a set standard Key Performance Indicators, taken from the Standard Nuclear Performance Model (SNPM), that corresponds to given processes and sub-processes.

4.2.10 Other important concepts

Business Intelligence is related to, or even consists of, a great spectrum of different technologies. In this section some of the more important concepts for this study have been described and some more common concepts like OLAP, OLTP, Data Mining and MIS, Dashboards, Scorecards and others are explained in *Appendix D*, but have been left out of the actual report, since it is not the main target for this study to explain Business Intelligence technically.

4.3 Nuclear Organisations

Different nuclear organisations around the world have played a great part in the development and the use of Key Performance Indicators. Here a brief description of the organisations most important for this study follows.

WANO – World Association of Nuclear Operators

The World Association of Nuclear Operators is a non-profit organisation created to improve safety at every nuclear power plant in the world. Set up in May 1989, after the accident at the Chernobyl nuclear power plant in 1986, to provide international cooperation to exchange operating experience, members can work together to achieve the highest possible standards of nuclear safety. It allows each operator to benefit and learn from others' experiences, challenges and best practice, with the ultimate goal of improving nuclear plant safety, reliability and performance levels, for the benefit of their customers throughout the world. Every organisation in the world, that operates a nuclear electricity generating plant, is a member of WANO. WANO neither have ties to commercial organisations nor any direct association with governments. [52]

NEI – Nuclear Energy Institute

The Nuclear Energy Institute is a twelve year old merge of a number of nuclear industry organisations forming this new nuclear interest organisation. "*NEI's objective is to ensure the formation of policies that promote the beneficial uses of nuclear energy and technologies in the United States and around the world*" [40]. NEI is a nuclear policy organisation and has, among other things, been one of the main initiators of the Standard Nuclear Performance Model, which is very important for this study. See further in *4.6 PAD framework*. NEI has 280 industrial members in 15 different countries. [40]

SKI – Statens Kärnkraftsinspektion

The SKI is the Swedish state authority, which controls all nuclear plants in Sweden. The SKI makes sure, that the holder of a licence to conduct nuclear activities in Sweden, fulfils the given responsibility to operate the facility safely. The SKI has an extensive cooperation with WANO. SKI's work is interesting for this study as it is one of the few organisations, which assess Swedish nuclear plants with performance indicators today. [48]

4.4 Future scenarios

The finding out of today's situation is an important part of this study, but the study also should form and measure possible system scenarios for the future. The scenarios are based upon the assumption, that the BI systems today, are more than one and that the information from the different systems are never joined together in a common DM or DW. Today's system situation will be referred to as s0.

s0. Uses the BI systems of the nuclear power plant today as basis. In particular three systems have been included in the study.

- Aktuell Prognos – Decision support system with one year forecasts
- SAP R/3 – Overall business management system.
- Fenix – Material maintenance system

There is to a great extent information exchange between the Material Management part of SAP R/3 and Fenix today and also information exchange between Aktuell Prognos and SAP R/3.

The following scenarios suggest three different ways of structuring the systems in the future.

s1. The first future scenario is very similar to what the systems look like today. The structures of the systems are kept and the modification work aims to achieve new KPIs, which are not used today, but can be found in the system without any structural changes. The work needed might include new database queries and use of new analytical techniques.

s2. The second future scenario is what can be called a DM scenario. The advantage of this scenario is gained by combining data from different BI systems, in order to create new information. The decision support information is still kept locally and joined together in a Data Mart. This gives new possibilities to cross process comparison and using data mining techniques to find new trends and dependencies in data.

s3. The third and last scenario is the full scale Business Intelligence structure, as shown above, where data from all systems goes to one enterprise unique DW joining information from the whole company. This gives unique possibilities to compare information from different units in the organisation. Greater amounts of data bring greater possibilities to find trends, see dependencies and predict future events. With clear interfaces, dependent DM are allowed to extract much more performance critical information about the whole enterprise directly from the DW.

The scenarios will be seen as possible futures for the BI systems and the questions in the frameworks will be asked, so that it becomes possible to determine which PAD a certain scenario would create, but as well as how hard it would be to get to the scenario measured in ME.

4.5 Frameworks

The EID frameworks in this study are based on standards, articles and documents by many prominent organisations and field experts, and they form the theoretical basis. The frameworks are also the base for all empirical research and the overall analysis. Both the PAD framework and the framework for ME are Extended Influence Diagrams based on Bayesian networks, which construct the overall analysis model, see also *6 Analysis*.

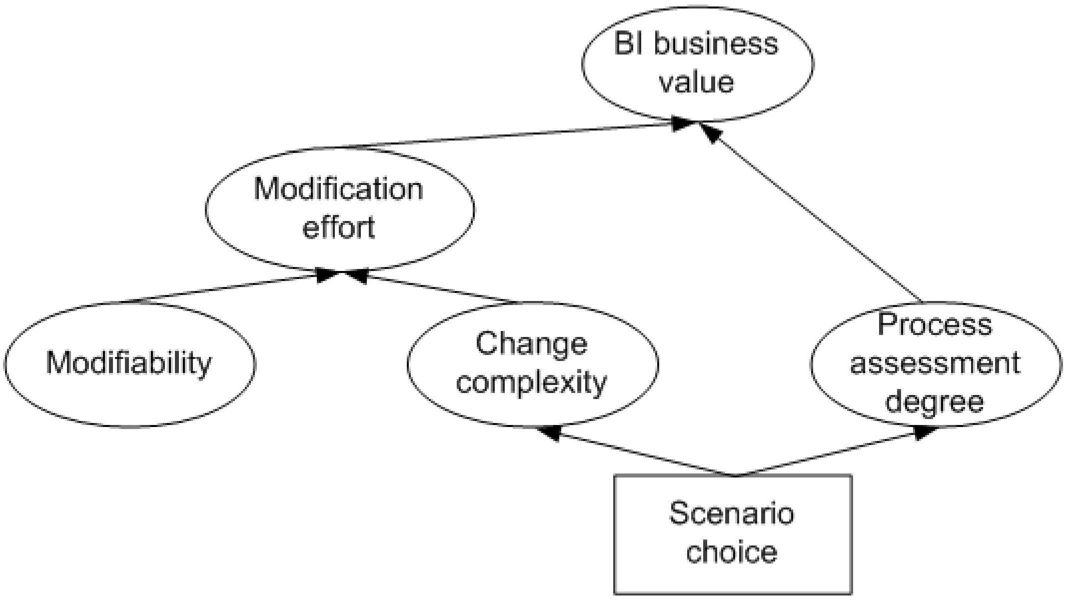


Figure 14, an overview of the theoretical framework.

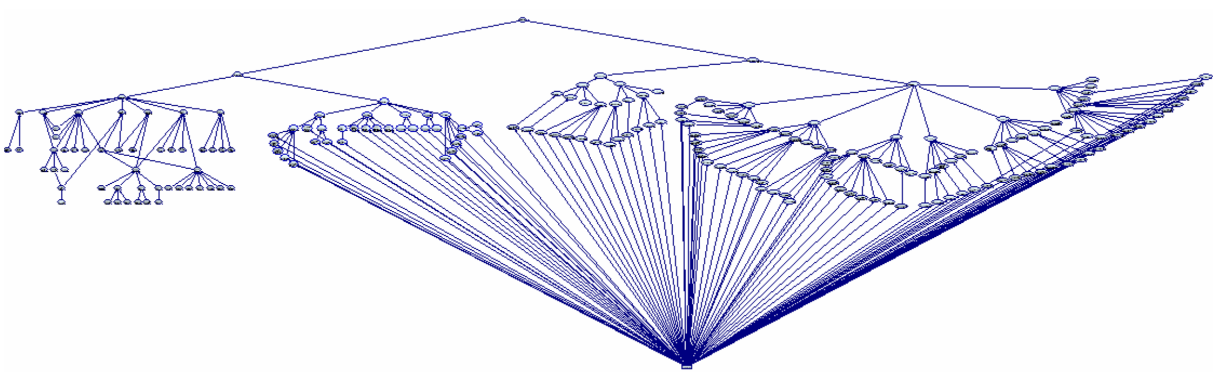


Figure 15, the whole theoretical framework.

4.6 PAD framework

The PAD framework consists of Processes, and Sub-processes. The KPIs all come from the SNPM [39]. The PAD framework is completely based upon this standard.

Some parts have though been cut out because of the following reason:

- All the processes concerning Human Recourses and Financing have not been included in the framework, since these fall outside the scope of this study. See *2.6 Delimitations* above.

The SNPM consists of processes and sub-processes, and to both of these there are linked KPIs, which assess their performance. Processes and sub-pocesses form two equally important parts of the PAD framework. The framework can simplified be described in the following picture.

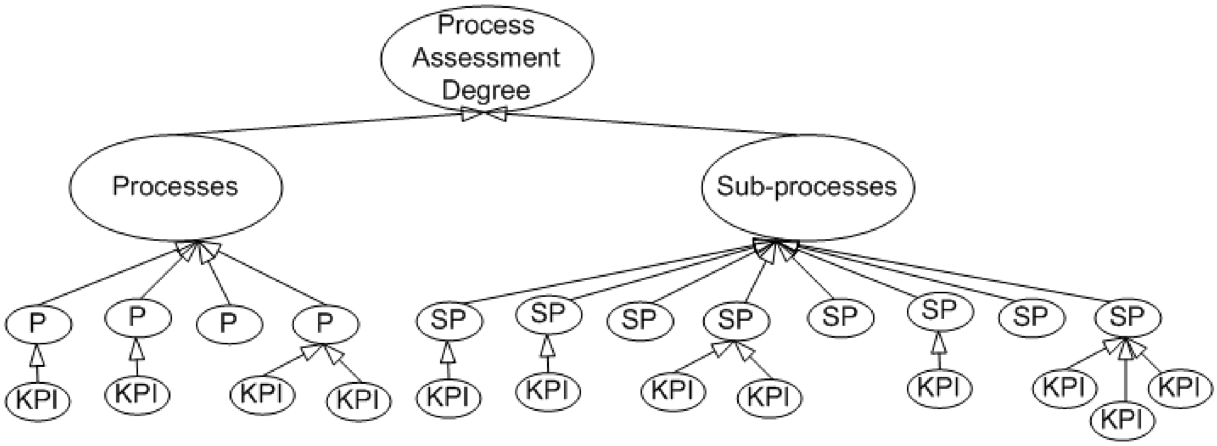


Figure 16, Schematic picture of the Process assessment degree framework.

In Figure 16 the processes and sub-processes form two independent sub-frameworks of the PAD framework. It is essential to understand that the sub-processes are not sub-nodes to the processes, but instead an independent part of the framework, at the same level as the processes. The reason for this is, that the KPIs of the sub-process normally do not assess the overlying process.

Processes and sub-processes with no attached KPIs to them, have not been measured. These processes and sub-processes have, however, been included in the prioritisation and they are still part of the study, even though it is uncertain how well these processes and sub-processes are assessed. The SNPM is a document, which is continuously developing, when more and more issues are agreed upon. The parts that lack KPIs today may easily be assessed in future work, if existing then, since the processes are already prioritised.

A complete description of all the processes and sub-processes and the KPIs in the framework can be found in *Appendix B*, but in order to communicate an understanding for them, a short example has been included in the main text as well:

EID reference	Name	Literature reference
SP37	WM008 – Monitor and Control Contamination	
Definition		
All activities associated with providing contamination control. Includes controlling and monitoring contaminated areas of plant, and providing decontamination services.		
Weight		
0,04500343		

Table 8, Example of a sub-process.

EID reference	Name			Literature reference
SP3KPI1	Effectiveness of Physical Change Authorization Processes (1)			
Definition				
of CRs written specifically against Physical Configuration change authorization processes				
State				Credibility
s0: -	s1: -	s2: -	s3: -	90%

Table 9, Example of a KPI.

4.7 Modification effort framework

The modification framework is not built on a single source, but is instead a joint theoretical base from almost 30 different sources as described in *3.4.1 Building the ME framework*. All the sources can be found in the end of the description of the EID in *Appendix A*.

The Modification Effort framework consists of two main parts, the modifiability and the Change complexity.

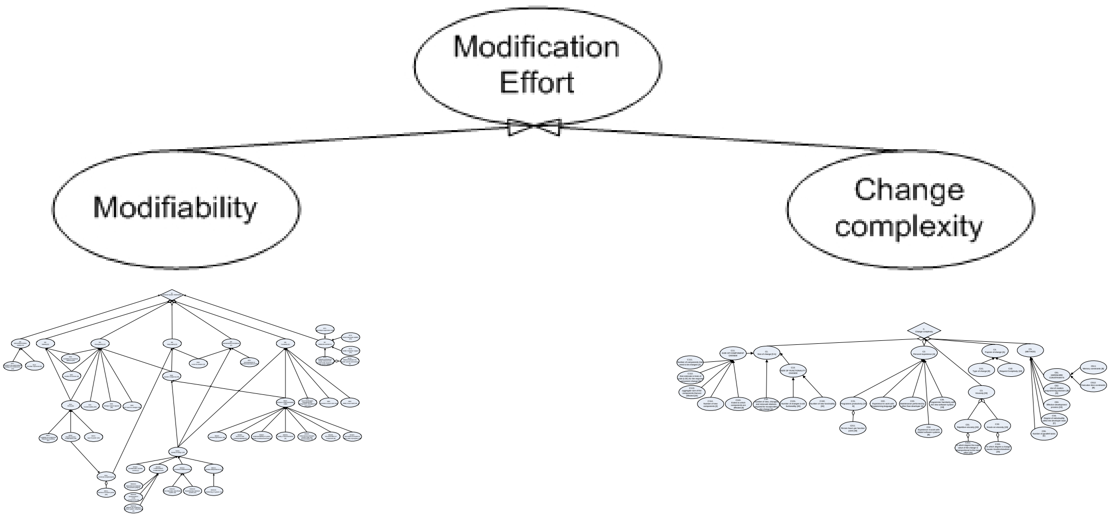


Figure 17, Modifiability and Change complexity form Modification Effort.

Modifiability and Change complexity are two completely independent EIDs. Together they construct the Modification Effort. Note that a high Modifiability lowers the Modification Effort, while Change complexity increases the value of the Modification Effort.

All the nodes have descriptions corresponding to the name and the node reference in *Appendix A*, but in order to communicate an understanding for them, a short example has been included in the main text as well:

EID reference	Name	Literature reference
M212	Number of Prior Modifications	[7]

Definition
Number of modifications made before. This can be measured in number of modifications requests, which have been carried out.

Question	State		
How many prior modifications have been performed on system #?	sys A: *L*	sys S: *L*	sys F: *H*

Criteria for analysis
If very many State = Low, else if medium State = Medium, else if almost none State = High.

Credibility			Weight
sys A: 90%	sys S: 90%	sys F: 90%	4.5

Table 10, Example of a node in the Modifiability part.

EID reference	Name	Literature reference
C32.	Knock-on Viscosity	[45]

Definition
To which Degree a Change Causes Trouble elsewhere

Question	State			
To which degree would the modifications needed for s# cause trouble elsewhere?	Person 1: s0: *L*	s1: *L*	s2: *L*	s3:
	Person 2: s0: *L*	s1: *L*	s2: *M*	s3:

Criteria for analysis
If lots of trouble State = High, else if some State = Medium, else no trouble elsewhere State = Low.

Credibility		Weight
Person 1 s0: 100%	s1: 60%	7
Person 2 s0: 100%	s1: 70%	

Table 11, Example of a node in the Change complexity part.

5 Empirics

As described in *3.5.1 Sources of evidence*, interviews are the most important source of evidence, but certain amount of documents have also contributed to a higher credibility and a better understanding of the systems included in the study. Great attention should be paid to the fact that the empiric data, for two of the three parts of framework, need to be collected for each of the four scenarios, in order to fully complete the analysis and in order to be able to give a scenario recommendation.

5.1 Interviews

In order to collect all necessary information about both the KPIs and the effort needed to do certain modifications, interviews with five persons have been made.

Interview 1, is concerning the PAD sub-framework and it included 3 persons. They were interviewed during one workshop-like interview, where focus was on which of the SNPM KPIs, that the nuclear power plant is using today and which could be used in the proposed future system scenario. During this interview the three main BI systems, for the SNPM processes and sub-processes, of the nuclear power plant were also determined.

Interview 2 & 3. are two separated interviews with two persons, concerning the ME framework. These interviews concerned the Modifiability of the three most important BI systems and about the Change complexity of the three future scenarios. The first interview, with a System Manager, concerned the two of the three systems: *Fenix* and *SAP*. The second, with another System Manager, concerned the third system, *Aktuell Prognos*.

In the table below, the interviewees are listed as well as the interview date. The questions and the answers for each interview can be found in the corresponding Appendix. All interviews were held at the nuclear power plant.

Interview	Interview person	Date	Corresponding appendix.
1	CFU	2006-11-28	*Appendix B*
1	System expert	2006-11-28	*Appendix B*
1	Financial manager	2006-11-28	*Appendix B*
2	System Manager 1 at the nuclear power plant	2006-12-07	*Appendix A*
3	System Manager 2 at the nuclear power plant	2006-12-07	*Appendix A*

Table 12, interview schedule.

5.1.1 Interview 1

During the first interview the focus was the KPIs in the SNPM. For each KPI included in the study, they were asked to answer, if the KPI be could put together at all, today. If the answer was yes, they were asked if the KPI could be put together manually or if a system put the KPI together.

Secondly, they were asked to say, if the KPI could be put together automatically, presuming simpler modifications were made to the system; like writing new queries to existing data bases, change of parameters or making simpler code modifications. Also if an automatic handling of the KPI needed added information from different local systems, combining data in a local DM. Or thirdly if they thought that the KPI could be put together only if data or information from outside systems were combined with information or data from local systems, preferably in a DW.

The first question determines the KPI state of s0 in accordance with the criteria of analysis in the table below. s1, s2 and s3 will be considered to have the same state as s0, if no change is suggested in the second question. The second question then determines the values of the future scenario as described in the table below.

When the first questions had been discussed the interviewees were asked to assess how sure they were on their answers. This created the credibility of the node together with the estimate made by the interviewer. In the following table an example of the answering sheet is shown. A sheet like this was used for each of the KPIs, in order to easily note the given answers during the interviews.

EID reference	Name	Literature reference
Definition		

Question	Answer
Do you identify and use this KPI and how is it put together?	○ We could not put together the KPI ○ It is put together manually ○ It is put together by a system ○ Don't know it
This KPI could be put together automatically if the systems you have today…	○ Would have simpler modifications ○ Collaborated more and saved information in a local Data Mart ○ Collaborated with systems outside of the nuclear power plant

Criteria for analysis

The first three options tell which Business Intelligence states, which exist today. " We could not put together the KPI " → State = *Low*, "It is put together manually" → State = Medium, "It is put together by system" → State = *High*

The second question addresses the future scenarios. If the first "Simpler modifications" is checked, all future scenarios will be given a PAD state *High*. If "Collaborated more and saved information in a local Data Mart" Future scenario 2 and 3 will be given PAD state *High*. Future scenario 1 will be given the s0 (the today scenario) state. Finally if "Collaborated with systems outside of the nuclear power plant" PAD state *High* will be given Future scenario 3 and for Future scenario 1 and 2 the PAD state will be the same as for s0.

In case no answer is given, the credibility is 0 % for the KPI.

Value	Göran experienced Credibility	Interviewee experienced Credibility	Weight

Table 13, question sheet for interview category 1.

The answers to the questions including the credibility assessment were given in consensus of the three interviewees of during this single interview occasion. More information about interview 1 including all the answers given is found in the *Appendix B*.

5.1.2 Interview 2 & 3

The interviews 2 & 3 consisted of two parts each. During the first part questions concerning the Modifiability in the BI systems at the nuclear power plant were asked. During the second, the questions were about the change complexity of future scenarios.

For the Modifiability part of the framework the base is the BI systems. Of the three most important BI systems, two were being questioned about during interview 2 and the third system during interview 3. For each system a question corresponding to one of the end nodes in the Modifiability part of the framework was asked. Depending on the node, the questions might be very different and so also the corresponding criteria for analysis. The answer might correspond to any of the three states for each of the systems.

Following the first question a credibility analysis was done. Note that credibility is assigned for each of the different systems. For Interview 1 it is not necessary to address the question about the importance of the node, since the weights in the PAD part of the framework are handled differently, but, as mentioned in *4 Theory*, the ME part of the framework is weighted by the interviewees during the interview sessions. Therefore the last question was how important the interviewee considers the particular node to be, for the Modifiability of a system in general. A node can either be *very important*, *quite important* or *very little important*.

Each of the BI systems is considered equally important. The answer and the credibility of each system affect the node just as much. The weight was determined by the importance estimated during the two interviews and the opinion in the two interviews was considered equally important to the weight of each node.

The following table is an example of what a question sheet looked like for a node. A and B in the sheet are systems in question, for example Fenix and SAP.

EID reference	Name	Literature reference
M11	Organizational Strategy for Maintenance	[3]

Definition

The long term plans concerning how the organisation should behave in order to get the level of maintenance efficiency as high as possible.

Question	Answer
Are there any written strategy documents for how maintainability of the system should be performed?	**A:** ______________________ **B:** ______________________

Criteria for analysis

If written strategy exists State = *High*, else State = *Low*.

Value			Credibility	**Interviewee**	**Me**	Weight
A: ***H***: ○	***M***: ○	***L***: ○		**A:**	**A:**	
B: ***H***: ○	***M***: ○	***L***: ○		**B:**	**B:**	

Figure 18, question sheet for first part of interview category 2.

The second part of the interview 2 & 3, concerning the Change complexity part of the framework, asks about the three future scenarios. It is very similar to the first part of these interviews and will be explained more briefly.

These interviews contain questions, for each future scenario and with all the BI systems in mind. In both interview 2 & 3 the interviewees have answered exactly the same questions and their answers are considered equally important for the state of the corresponding node.

Credibility is determined in the normal way for each the scenarios and weight has been determined in the same way as for the first part of these interviews. The following table is an example of a question sheet used for each node of the Change complexity part of the framework.

EID reference	Name	Literature reference
C22.	Experience in Work with Programming Language	[36]

Definition

The degree to which the programmer / integrator has previous knowledge of the programming language.

Question	Answer
How experienced are the programmers to work with changes in the systems with changes like scenario …	Scenario 1: ____________________ Scenario 2: ____________________ Scenario 3: ____________________

Criteria for analysis

If very experienced State = *Low*, else if medium State = *Medium* else little experienced State = *High*.

Value			Credibility	Interviewee	Me	Weight
S1: ***H*****:** ○	***M*****:** ○	***L*****:** ○		**s1:**	**s1:**	
S2: ***H*****:** ○	***M*****:** ○	***L*****:** ○		**s2:**	**s2:**	
S3: ***H*****:** ○	***M*****:** ○	***L*****:** ○		**s3:**	**s3:**	

Figure 19, question sheet for second part of interview category 2.

5.2 Documents planned to be part of the empirics

In order to prepare for the interviews, as much documentation about the programs and the structure of the IT systems as possible, was asked for. More specific documentation [13, 14] about *Aktuell Prognos* have also been used to strengthen the result by confirming the answers to questions corresponding to some nodes in the Modifiability part of the framework and thereby enhance the validity of the study.[54]

6 Analysis

The empiric data was analysed with the theoretical frameworks as a basis and in accordance with methods for analysis earlier discussed. In this chapter a presentation on how these methods were applied on each of the different framework parts will be given. An explanation is given to how results were produced and how problems with, for example, lack of data were handled. Furthermore, an explanation on how the different parts of the framework are connected together will be presented, which is important for understanding the result.

6.1 Practical ways of working and analysis-aiding software

The analysis was generally done in accordance with the methods described in *3.6 Method for the Analysis* and was greatly aided by a few software programs. This short chapter aims to describe how the analysis was performed practically.

The answers given during the interviews were saved as tables of data in Excel [35]. The data was manipulated to fit the program used for the mathematical analysis, called GeNIe [10]. In GeNIe it is possible to represent EIDs graphically and assign values to the end nodes. GeNIe also provides the possibility to introduce scenarios with the same mathematics as used by the EIDs [10]. This means that the work of aggregating values was done automatically.

The data from the interviews were entered in GeNIe as probability tables for the end nodes. In order to aggregate values to the top, GeNIe also needs dependency tables which show how each state of every parent affects the child node. The dependency tables can become very large, so in order to create these tables, weight data was exported from Excel to Matlab [33], which is a software for operating advanced mathematical functions and for presenting results efficiently. A Matlab function was written to create the tables and then to export them into text files, from which they could easily be copied into GeNIe. Some of the dependency tables were set manually on the higher levels to form the special constellation between the different parts of the theoretical framework. See further *6.2 The framework connections*.

6.2 The framework connections

The theoretical framework consists, as could be seen in chapter *0,* of two main parts. The most important task for the analysis is to set causal relations and the dependency tables between these two parts so that they correspond with the underlying idea of the theoretical framework. The dependency table between the two parts of ME, Modifiability and Change complexity, are also very important and will be described here.

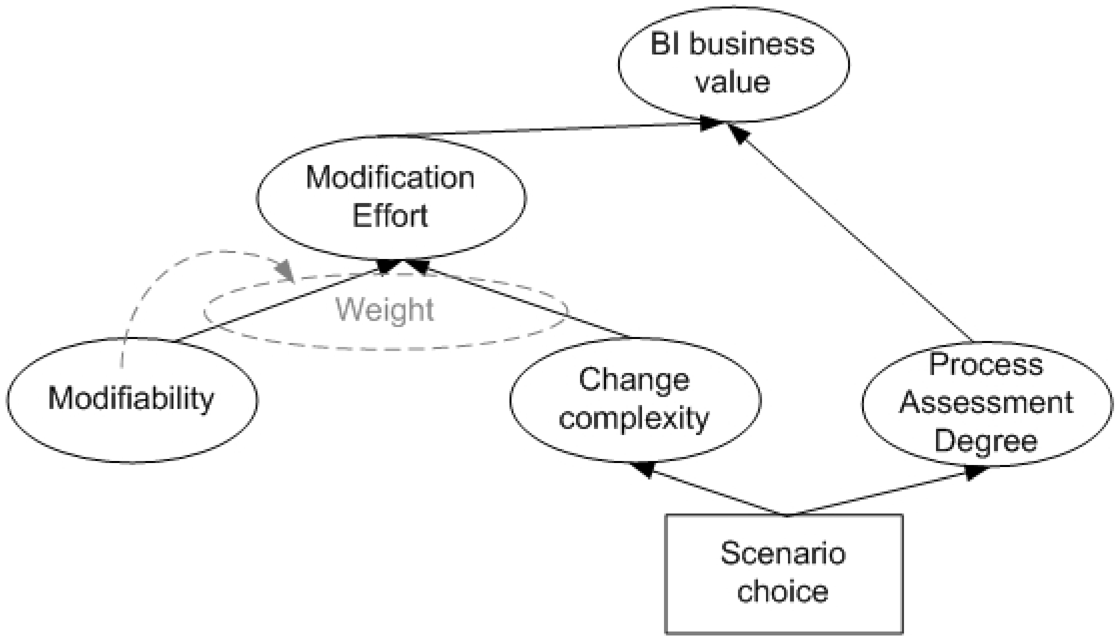


Figure 20, the three main parts; Modifiability, Change Complexity and Process Assessment Degree. The Modifiability also influences the weight between Modifiability and Change complexity.

The three interview parts; Modifiability, Change Complexity and Process Assessment Degree are all important for the BI business value which is the utility node, on top of the framework.

In the cost part, the weight between Modifiability and Change complexity depend on the State of the Modifiability. If the Modifiability is *High* the influence of Change complexity is lower than if it is *Low*. This hands-on change to the framework is necessary to maintain the original idea of ME. Without letting the influence of Change complexity depend on the value of Modifiability the ME would simply be the added sum of its two parents. With this solution the ME becomes more of a product of its parents which is exactly what the ME should be.

The dependency table of ME is therefore set from the non-scenario-dependent sub-framework of Modifiability. This can be done by considering Modifiability utility node of a sub-framework and the *Uncertain* state probability removed by normalising the probabilities over only the three remaining states. In order to produce a value which can be used to create the dependency table of ME, the three states are then assigned the following values: *High* = 50, *Medium* = 0, *Low* = -50. The value for the utility node = *M*.

$M / 100 = m$

m then influences the dependency table the following way. Note that the Modifiability affects ME positively and the Change complexity affects ME negatively:

Mod.	High				Medium				Low				Uncertain			
Change complexity	H	M	L	U	H	M	L	U	H	M	L	U	H	M	L	U
High	0.5 - m	0	0	0	0.5 - m	0	0	0	1	0.5 + m	0.5 + m	0.5 + m	0.5 - m	0	0	0
Medium	0	0.5 - m	0	0	0.5 + m	1	0.5 + m	0.5 + m	0	0.5 - m	0	0	0	0.5 - m	0	0
Low	0.5 +m	0.5 + m	1	0.5 + m	0	0	0.5 - m	0	0	0	0.5 - m	0	0	0	0.5 - m	0
Uncertain	0	0	0	0.5 - m	0	0	0	0.5 - m	0	0	0	0.5 - m	0.5 + m	0.5 + m	0.5 + m	1

Table 14, dependency table of BI business value.

In this study, *m* was calculated to *m* = 0.1435175, following the results of Modifiability, see further in *7.6 Modifiability*.

A problem that could affect the result gravely is the uncertainty. Especially if the uncertainty of the cost and benefit parts of the framework are very different in size. This is likely to happen in this study since the benefit part of the framework by default has a 32 % uncertainty caused by the incompleteness of the standard upon which it is built. In practice this could mean that even though 100 % of the given answers on the benefit side point toward a *High,* only 68 % will be affecting the BI business value. This would not cause any difference to the result if the uncertainty were equal in size on the two sides, but if not, it works just the same as a weight system where the uncertain - or in this case uncompleted - side of the framework would suffer. Therefore a complementary result will be calculated without uncertainty in the parents (ME and PAD) of the top node (BI business value).

6.3 Process Assessment Degree framework analysis

The PAD part of the framework is a bit different from the rest. This is mainly because it has been built up from one fixed standard, rather than many various sources. This part of the framework has two main differences when it comes to the analysis; the weights, and the great lack of measurable end nodes. Before looking into the uncertainty of the PAD part, a short update on this parts constellation and the data collected for this part will be given.

The PAD part of the framework consists of processes and sub-processes assumed equally important for the Process Assessment Degree. They may have one or more KPIs attached to them, which can be used for assessing the assessment degree. The empiric work focused on which KPIs that are used today and which KPIs could be used in the future, if systems were changed in accordance with the scenarios.

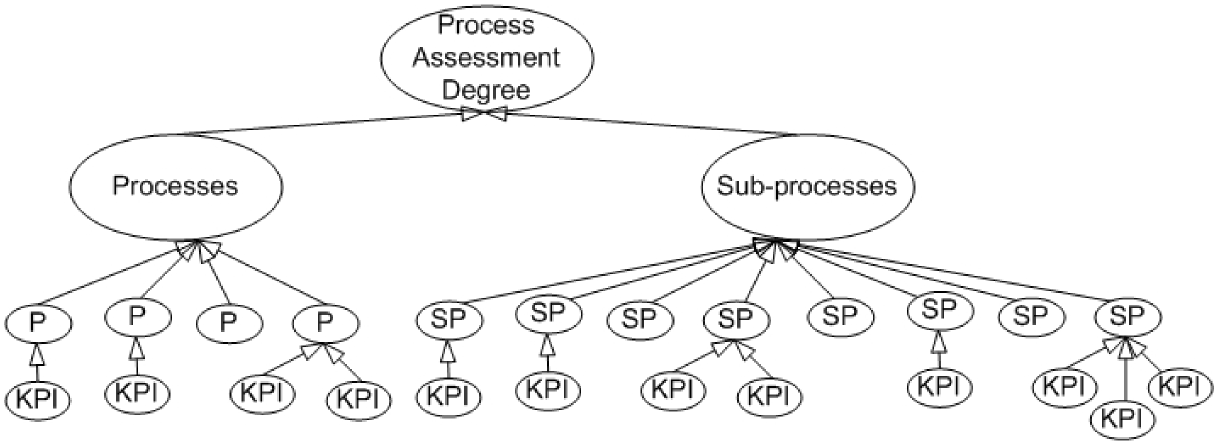


Figure 21, Schematic picture of the PAD part of the framework.

Data for the PAD was, as can be seen in *5 Empirics*, collected during one single interview with three persons. The weighting of this part of the framework was done before the interview by the BIO for the utility company. Each node of measurement has always only one source of evidence since the three persons interviewed came to a consensus on each question.

The input data for the analysis are, for each scenario, which KPIs could be put together automatically, which can be put together manually, given the existing information, and which ones are not possible to put together at all. The credibility estimation has been performed as it is described in *3.6 Method for the Analysis*. The weights have been set on the processes and sub-processes separately so that each process has a nominal weight towards all the other processes and all the sub-processes have a nominal weight towards all the other sub-processes.

The way of weighting the non-end nodes is different from the rest of the framework and from methods described in *3.6 Method for the Analysis*. In the case that the process or sub-process has more than one KPI, the KPIs are assumed equally important, with equal weight.

For many of the processes and sub-processes the SNPM does not recognise any KPIs. The standard is under continuous development and these processes and sub-processes may have many KPIs in the future, even though they have not been agreed upon yet. In order to let the PAD part of the framework reflect the SNPM as much as possible, no processes and sub-processes have been left out. Instead, the immeasurable KPIs are given a 100 % probability to be uncertain. This of course creates a great uncertainty in this part of the framework since 25 % of the processes and 39 % of the sub-process are not assessed.

The sub-process part consists of 38 sub-processes which all connect to the same child. This creates some practical problems since this brought about a dependency table that was a size of $4 * 4^{38}$. These tables take GeNIe way to long to handle. In practice a dependency table for a child having 8 parents takes 25 minutes to paste into GeNIe. Therefore 8 was set as the maximum number of parents for this study. The sub-processes were therefore divided into six groups, each having an intermediate level child node which was then connected to the top level Sub-processes node. Processes were divided into two groups the same way. Note that this did not affect the importance of the sub-processes towards the top level Sub-processes node. Values in the dependency tables in the intermediate nodes and in the top node have been recalculated so that the original weights have been preserved.

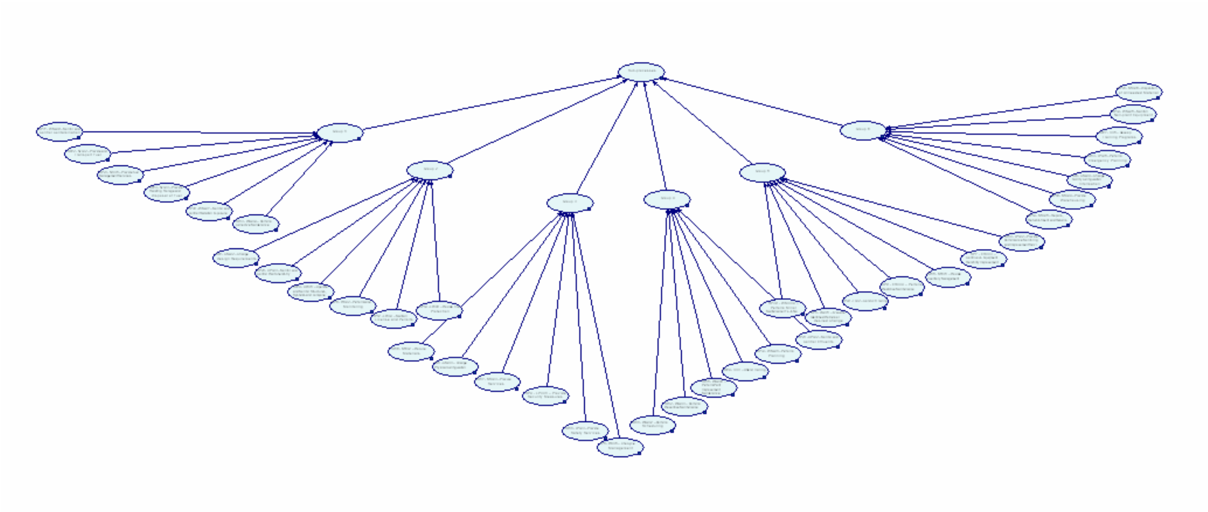


Figure 22, the sub-processes are divided into six groups with a intermediate node per group

6.4 Modifiability framework analysis

Modifiability is the first of the two parts that together form the Modification Effort part of the framework. What makes the modifiability part different from the rest is that it is not dependent on the scenarios. Instead Modifiability is build upon the attributes of the Business Intelligence systems of the nuclear power plant.

The Modifiability framework is the most nestled and least tree-structured of the framework parts. This does not mean any difference for the analysis but demands that extra good care is taken when calculating the weights and the probabilities. For any given child each of its parents carries a weight that determines how much it influences the child compared to the other parents, i.e., the weight of such a group of parents sum up to one.

Data has been collected during two interviews, where the first covered two of the three studied BI systems, and the second the third system. Generally each end node gets a third of its state probabilities from each of the systems. The three systems are independent of each other, implying that they have often shown very different modifiability when it comes to different aspects of the framework. Therefore, the typical modifiability end node probabilities are better distributed than in other parts of the framework. In the case where a question could not be answered for one of the systems, the state was 100% *Uncertain* for this system's share of the total state probabilities.

Credibility estimations have, in accordance with what is described in *3.6 Method for the Analysis,* been applied to the interviews and the visual proof from each of the three systems. This means that credibility for each end node, in the case of interviews being the only evidence, consists of 6 estimates. Visual proof has been used to strengthen the answers and the credibility of the analysis in some cases.

The weight system in this part of the framework works in accordance with what is described in *3.6 Method for the Analysis*. Both of the interviewees have estimated how important each end node is for the Modifiability on a three level scale (*Very little importance*, *Quite important* and *Very important*). These levels have been translated into numbers so that the weights can be calculated. The translation has been done together with Jonas Öhrström and Robert Lagerström both studying Modifiability for the Royal Institute of Technology.

Grade of importance (linguistic) in interviews	Grade of importance (numeric) in analysis
Very little importance	1
Quite important	4
Very important	10

Table 15, translation table from importance grade as answered in interviews and as calculated as weight within analysis.

The table should be understood as that if the interviewee answered *Quite important* it means 4 times as important as if the answer were *Very little importance* and 2.5 times less important than something which is *Very important*.

The importance of each end node has then been calculated as an average of the two answers given. Nodes are compared within parent-with-one-common-child groups.

6.5 Change complexity framework analysis

The Change complexity is the second sub-part of the Modification Effort part of the framework. Just as in the PAD part the Change complexity is dependent on scenarios. Where PAD can be seen as the beneficial part of each scenario and the Change complexity is the cost part of each scenario.

Two interviews asking the same questions have been made with two persons for each end node and for each scenario. Determining on how laborious each scenario would be to achieve from the nuclear power plant´s perspective on each one of the end node perspectives.

Credibility and Weight has been produced the same way as in *6.4 Modifiability framework analysis*. The important difference is that Credibility has a new additional impact on the probabilities. Since the questions in this part of the framework have been asked independently and with the same basis to two different persons, the answers can be seen as complements to each other. Therefore the answer with the higher credibility has been given slightly higher impact to the probability table of the corresponding node. The 1 – 3 grading scale used for estimating credibility is used; see *3.6.1 Probabilities of end nodes and credibility of sources of evidence*. An answer with average estimated credibility 3 will count 1.5 times more than an answer with average estimated credibility 2.

This is best shown in this short example:

Person 1 answers a question with *High* and with an average estimated credibility of 1.

Person 2 answers a question with *Low* and with an average estimated credibility of 3.

The total credibility is then (0.9 + 0.5)/2 = 0.7

Person 2's answer counts 3 times as much as Person 1's.

Low = 0.7 * 0.75 = 0.525

High = 0.7 * 0.25 = 0.175

We get/This yields the following probability table:

Example node	
High	0.175
Medium	0
Low	0.525
Uncertain	0.3

Table 16, the probability table.

6.6 The result presentation

In the top of the framework is the BI business value, and for this node, as for all nodes, the state probabilities will be calculated. The state probabilities form the basis for the results of the study. These probabilities are shown in circle diagrams for each of the scenarios. This can also be done for sub-parts of the framework.

The probability circle diagrams are not optimal for scenario comparison, so instead an easily comparable value quantifying BI business value of each scenario is calculated. This is not only done for the main framework but also for each of the sub-parts of the framework. In GeNIe is this done by the introduction of the utility node.

To create a value scale reaching from 0 to 100 the following values are assigned to each state:

	High	Medium	Low	Uncertain
Value	100	50	0	50

Table 17, the different states get different values assigned to them in order to get the values of the node.

The result from multiplying the probabilities of each state, in any given node, with the corresponding value from *Table 17,* the different states get different values assigned to them in order to get the values of the node. is a approximated value, for the node, ranging between 0 and 100. The value 100 means that there is a 100 % probability that the state of the node is *High*. The different values of a node for each of the scenario can then be easily compared in a bar diagram.

By calculating the value of the node with the probability of *Uncertain,* as if it was first *High* and then *Low,* two different values - denoting the spread of uncertainty - are obtained. This is shown as a black line attached to the top of each staple in the diagram. Theoretically, the value of node could be anywhere on this black line.

7 Results

All results have been produced in accordance with *3.6 Method for the Analysis* and *6 Analysis* and strive to deliver a correct answer to the question of the study.

7.1 First important result of the study: The framework

The theoretical framework with its subparts is not only a part of the analysis leading to the results but it is also the answer to the sub-questions for this study:

1. *How can the Process Assessment Degree be assessed?*
2. *How can Modification Effort be assessed?*

The work with the frameworks themselves, the weights, and the analysis methods connected to them are very important results on how PAD and ME can be assessed. The frameworks, the description of the nodes, and the weights connecting them are described in *Appendix A and Appendix B.*

7.2 BI business value and Choice of scenario

When summing up the parts of the framework it leads to the final result; the BI business value.

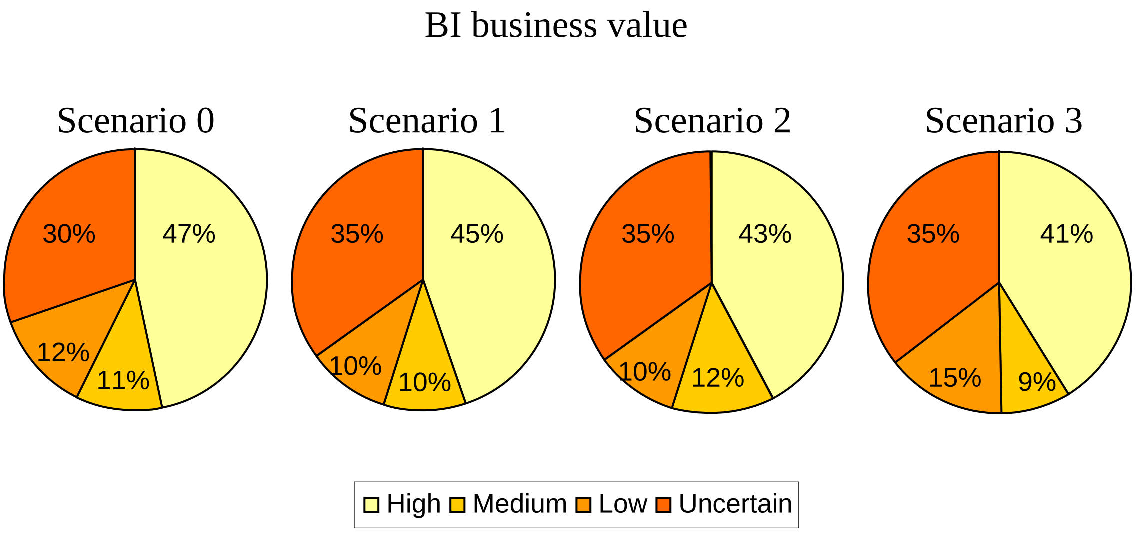


Figure 23, BI business value state probabilities for scenarios 0 through 3. This diagram shows four very similar diagrams.

What is seen in the diagram above and what becomes even clearer in the corresponding value diagram below is the how hard it is to tell which scenario is the better.

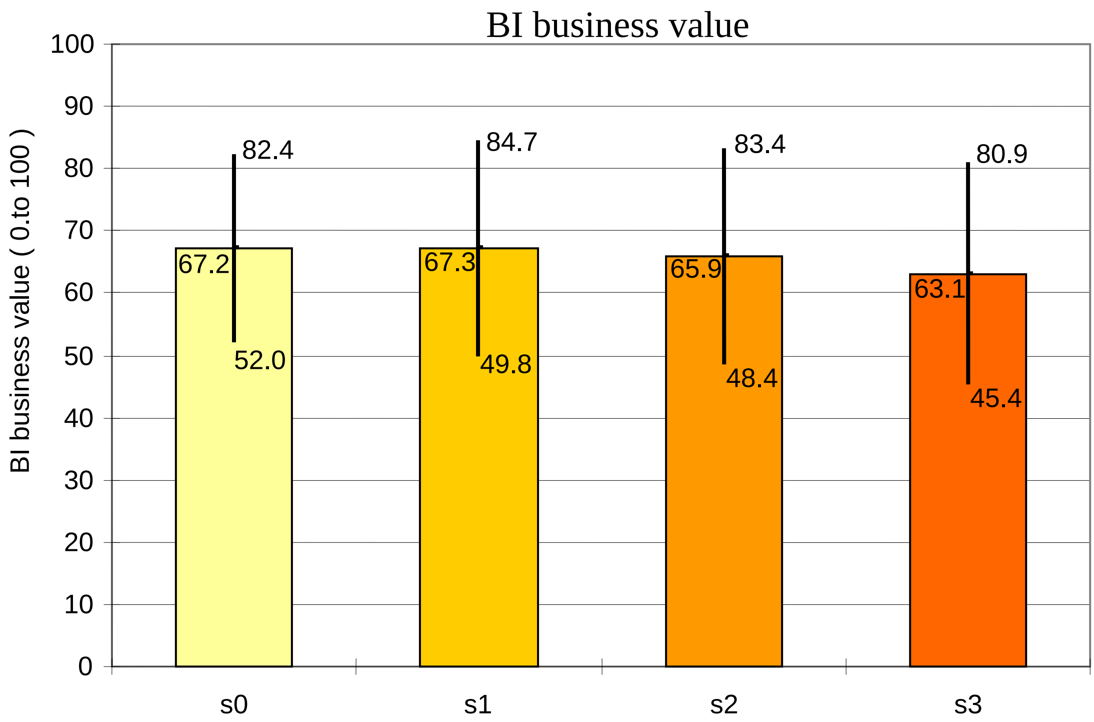


Figure 24, value diagram for BI business value and the spread of uncertainty.

The great uncertainty and the high similarity of values make it impossible to recommend any of the scenarios simply from this study. Just the fact that scenario s0 (no changes) has such a high value makes it questionable whether any change at all should be implemented with basis on this study.

In accordance with what was discussed in *6.2 The framework connections,* the BI business value will also be calculated when the parents' uncertainties are eliminated. This provides only a slightly different picture but shows that the non-balanced parents' uncertainties' affect the result:

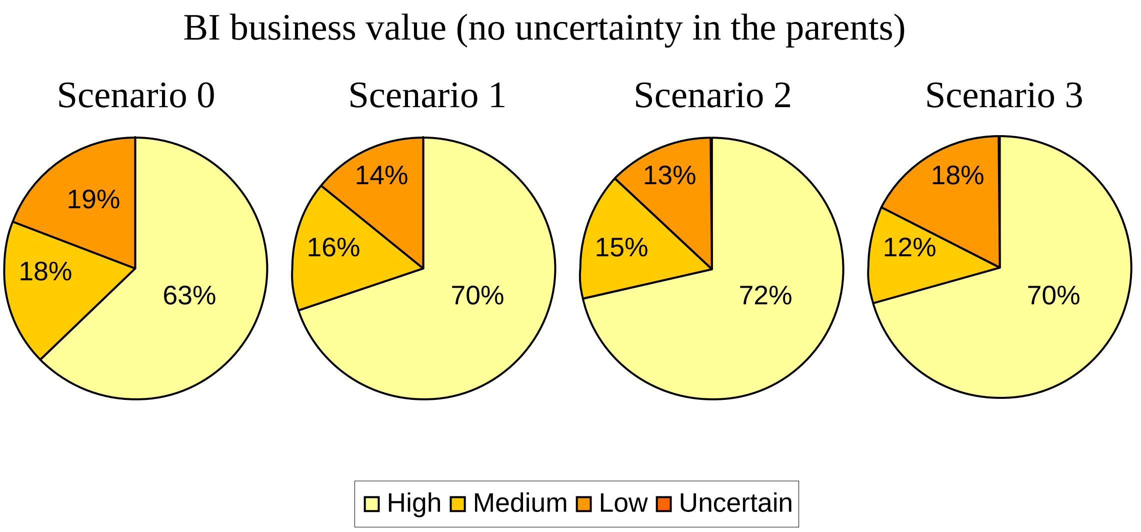


Figure 25, the circle diagrams of BI business value when the uncertainties have been taken away from the parents.

In the following value diagram, based on the circle diagrams above, scenario 1 and 2 have the strongest numbers:

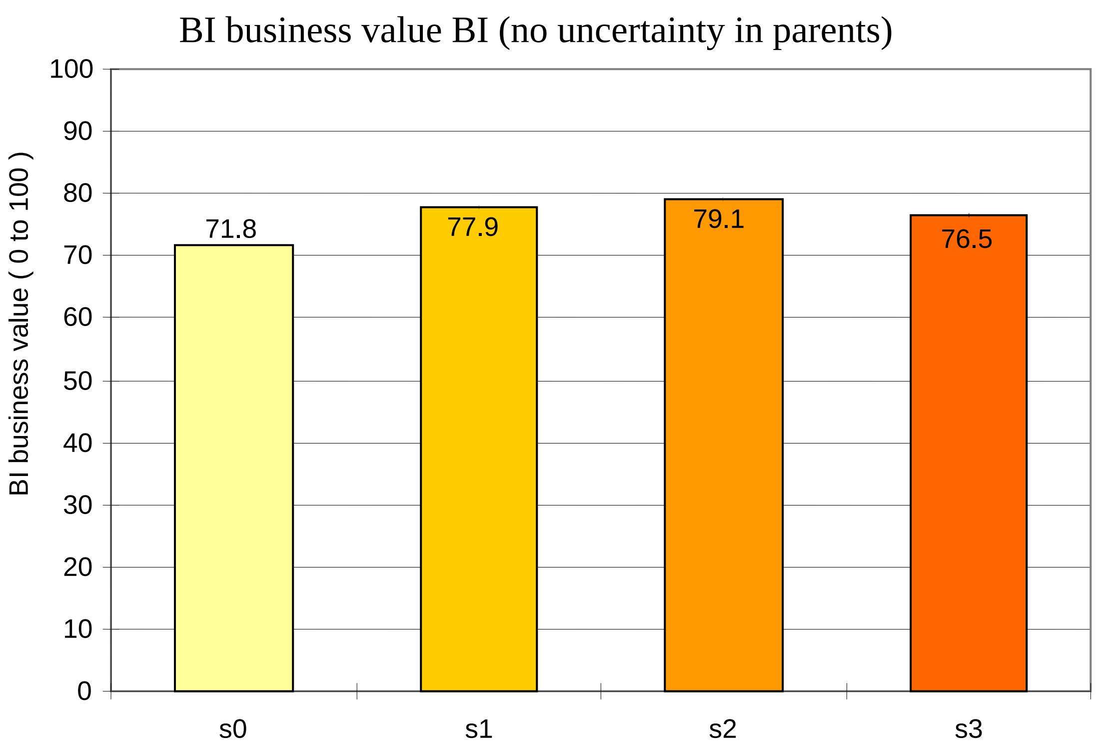


Figure 26, value diagram of BI business value, calculated without uncertainties in the PAD and ME.

Neither this result can be used for making decisions about which scenario really is the better, since all scenarios have so similar values, but this result is probably a bit closer to the truth when compared to what could be seen in *Figure 24*.

7.3 Answer to the question

From the results presented an answer to the question of the study will be answered:

Which future system scenario enables the highest Business Intelligence (BI) business value in the Decision Support Systems of the nuclear power plant?

The answer is:

No system scenario has a higher ratio then another. No system scenario can be recommended!

Due to the high uncertainties and the very similar BI business value between all scenarios no scenario has proved to be better than the other. This makes it impossible to recommend any of the scenarios based on the result of this studies quantitative theoretical framework results.

A recommendation should instead be made on other attributes of the offered system solutions. See a discussion about these attributes in *8.4 Other attributes of BI systems to consider*.

7.4 *Process Assessment Degree*

The nuclear power plant only has systems that put together 35 % of the 69 KPIs asked about during the interviews. The future scenarios could provide a higher PAD. The following circle diagram shows the PAD result.

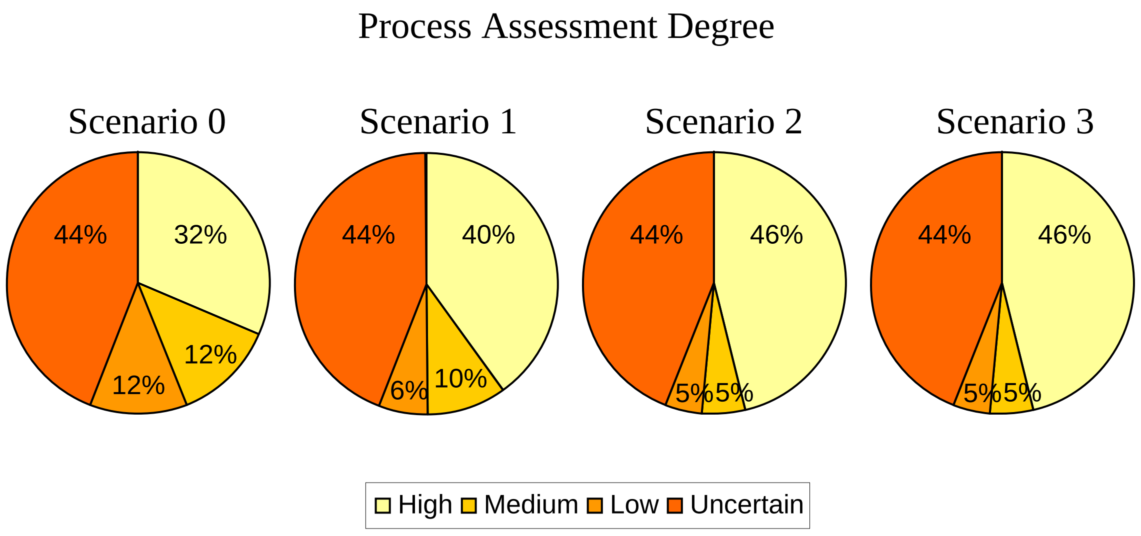


Figure 27, Process Assessment Degree state probabilities for scenarios 0 through 3.

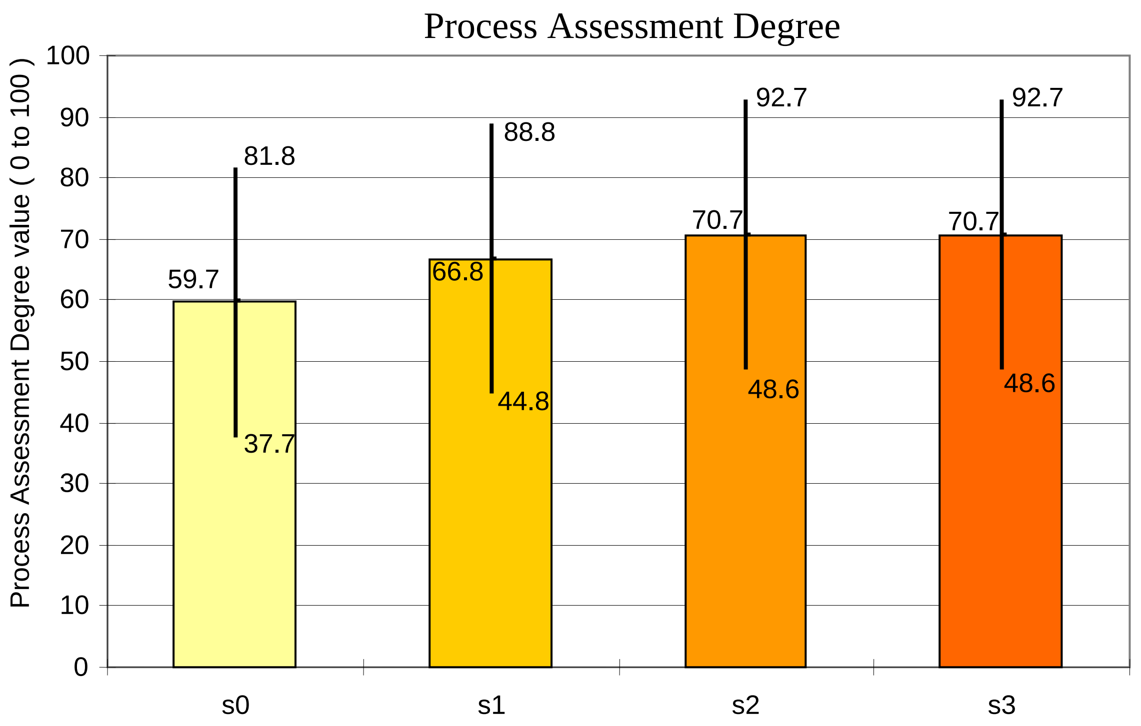


Figure 28, value diagram for Process Assessment Degree and the spreads of uncertainty.

7.5 Modification Effort

ME is the product of Modifiability and Change complexity. The result is shown in the circle diagram below.

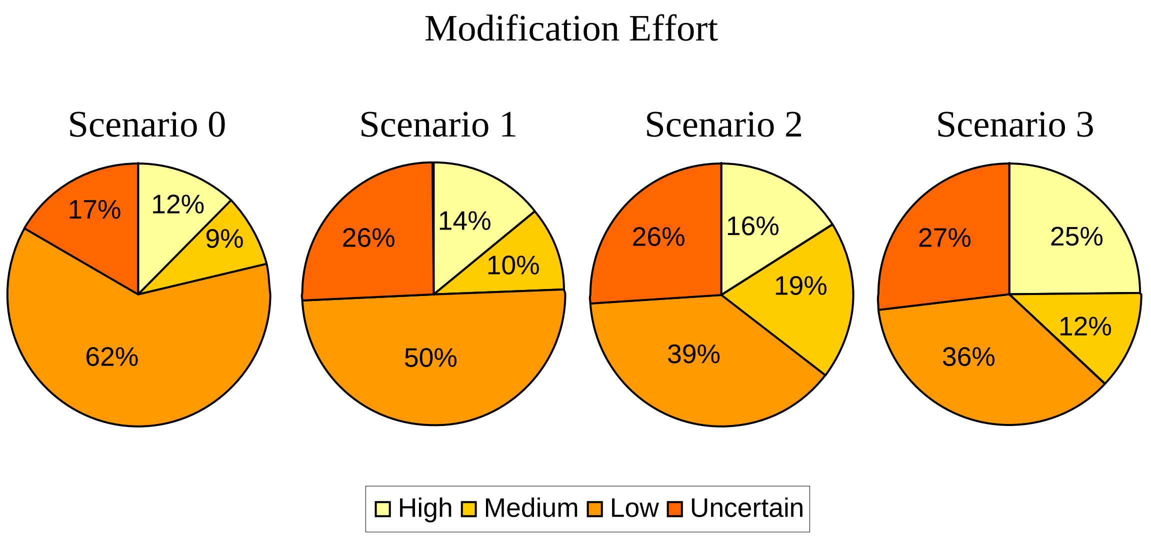


Figure 29, Modification Effort state probabilities for scenarios 0 through 3.

Here follows a value diagram of ME:

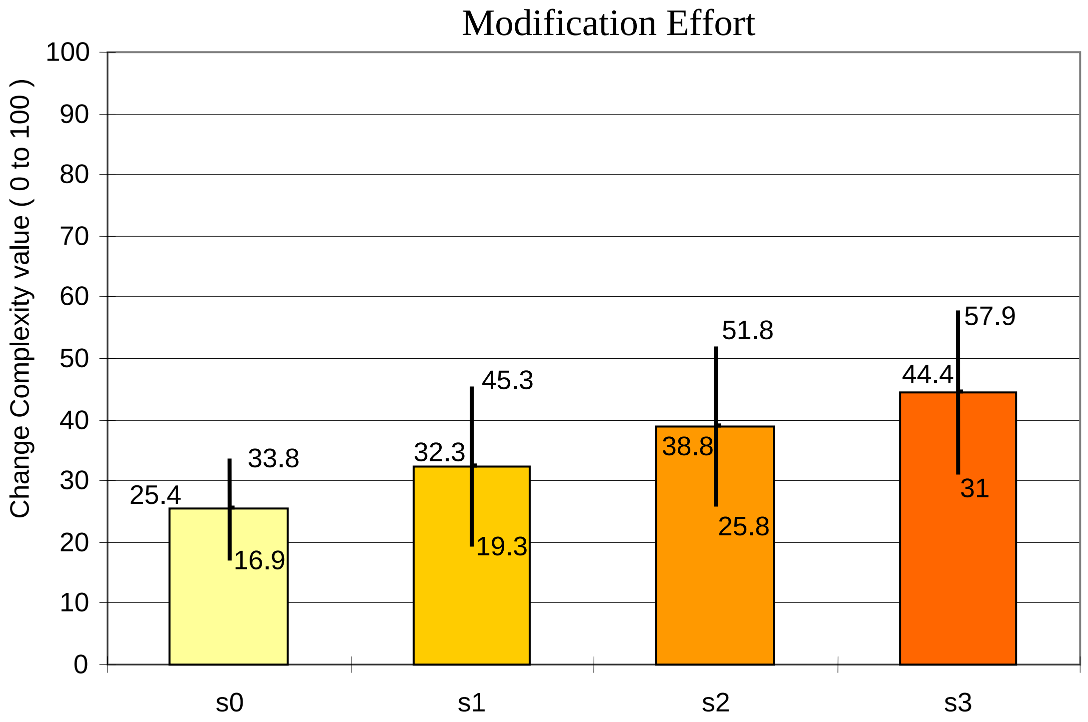
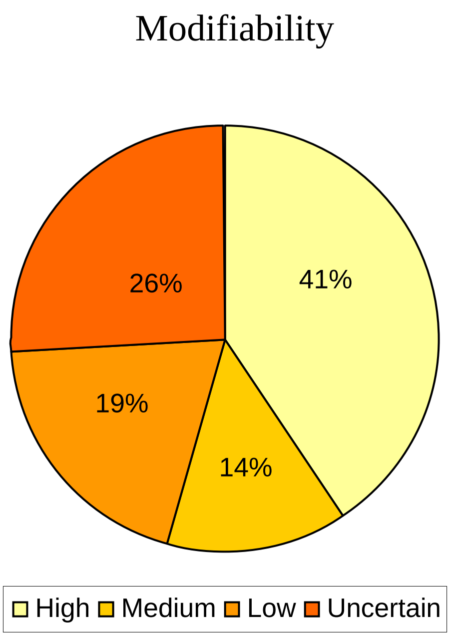


Figure 30, value diagram for the Modification Effort and the spreads of uncertainty.

7.6 Modifiability

As can be seen in the circle diagram below the probability that the Modifiability is *High* is 41 %, while *Medium* and *Low* only have a probability of 14 % respectively 19 %. Even though the uncertainty is 26 % it is probable that the modifiability of the BI systems of the nuclear power plant is *High*.



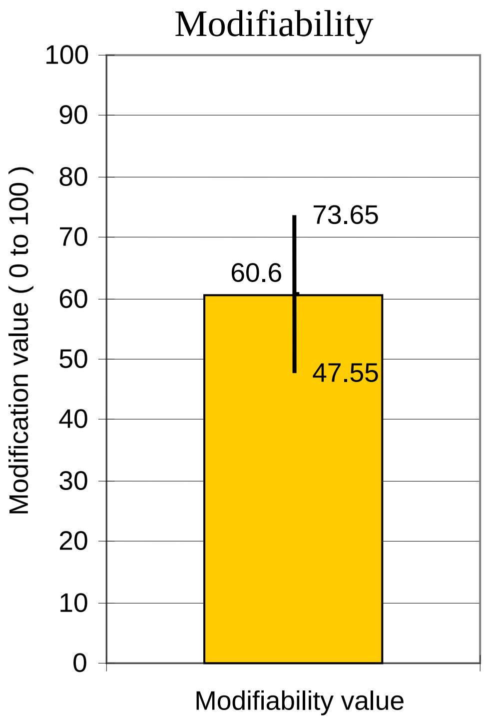


Figure 31, probabilities of the different states for Modifiability.

Figure 32, the value of Modifiability and the spread of uncertainty.

In order create the dependency table for ME it is necessary to calculate the value *m* as described in *6.2 The framework connections*. Following the calculations described in chapter *6.2* it is found that *m* = 0.1435175.

7.7 Change complexity

In the following circle diagrams is it possible to see a clear development from left to right through the different scenarios.

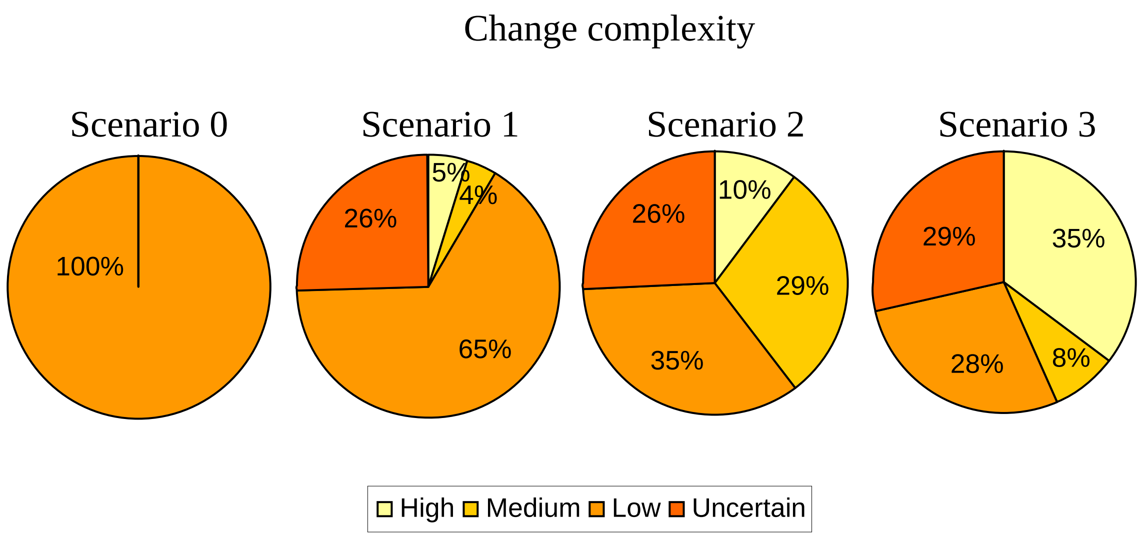


Figure 33, state probabilities for Change complexity in scenarios 0 through 3.

Following the probabilities the value diagram for Change complexity is obtained.

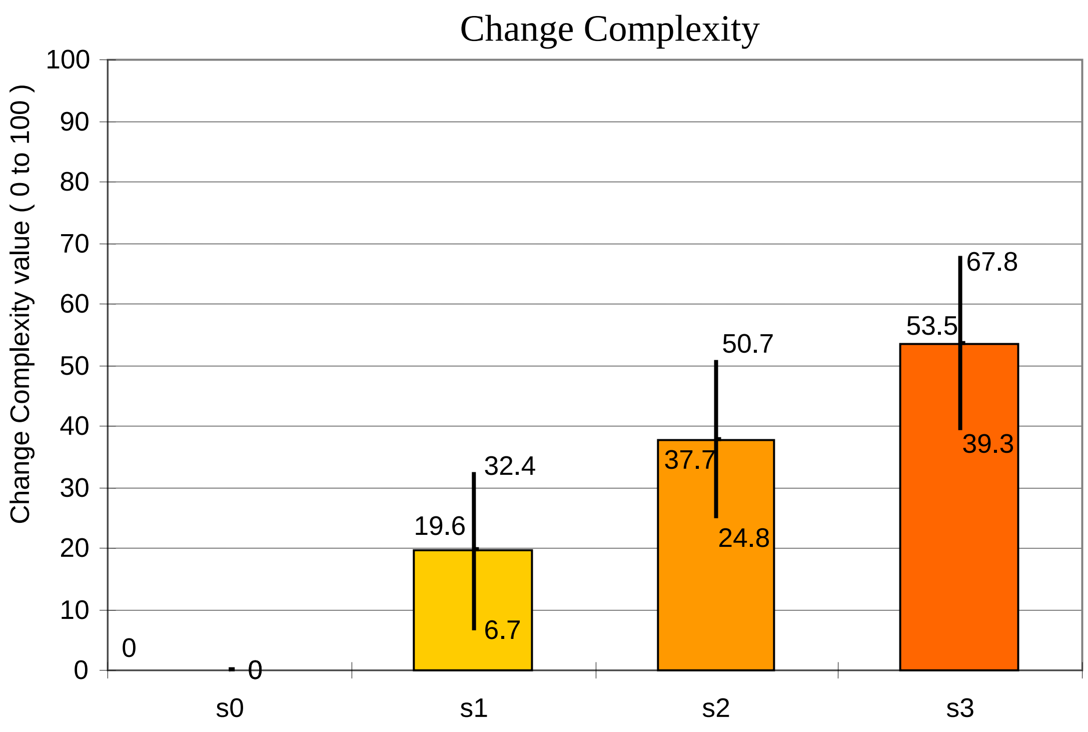


Figure 34, the values of Change Complexity and the spread of uncertainty.

8 Discussion

In this chapter the result itself is discussed, for example, why no scenario recommendation could be given. Aspects of the result that have not yet been covered are also discussed.A discussion on the credibility of the result is covered and an analysis of the sources of uncertainty is done. The empiric research is reviewed and, finally, improvement potentials in the BI systems at the nuclear power plant are identified.

8.1 Result discussion

When having set out to recommend one out of three future scenarios, it is of course not pleasant to not be able to do so. The possibility that the three scenarios would end up with such similar values, and that uncertainty would make it impossible to recommend any of them is, however, larger than one may first think. Considering certain aspects of this study it is in fact very likely that the results end up in something where no recommendation can be given.

The first of those aspects that needs to be considered is the fact that the study is a cost/benefit analysis. The scenarios reflect possible futures of reality and they are not surprisingly of such kind that a lower cost normally is associated with a lower benefit. When looking into the results on both sides of the analysis, it is found to be exactly so. The diagram below shows how the cost and the benefit value curves are inverted towards each other. The higher the scenario, the higher the costs but also the benefits that come with it.

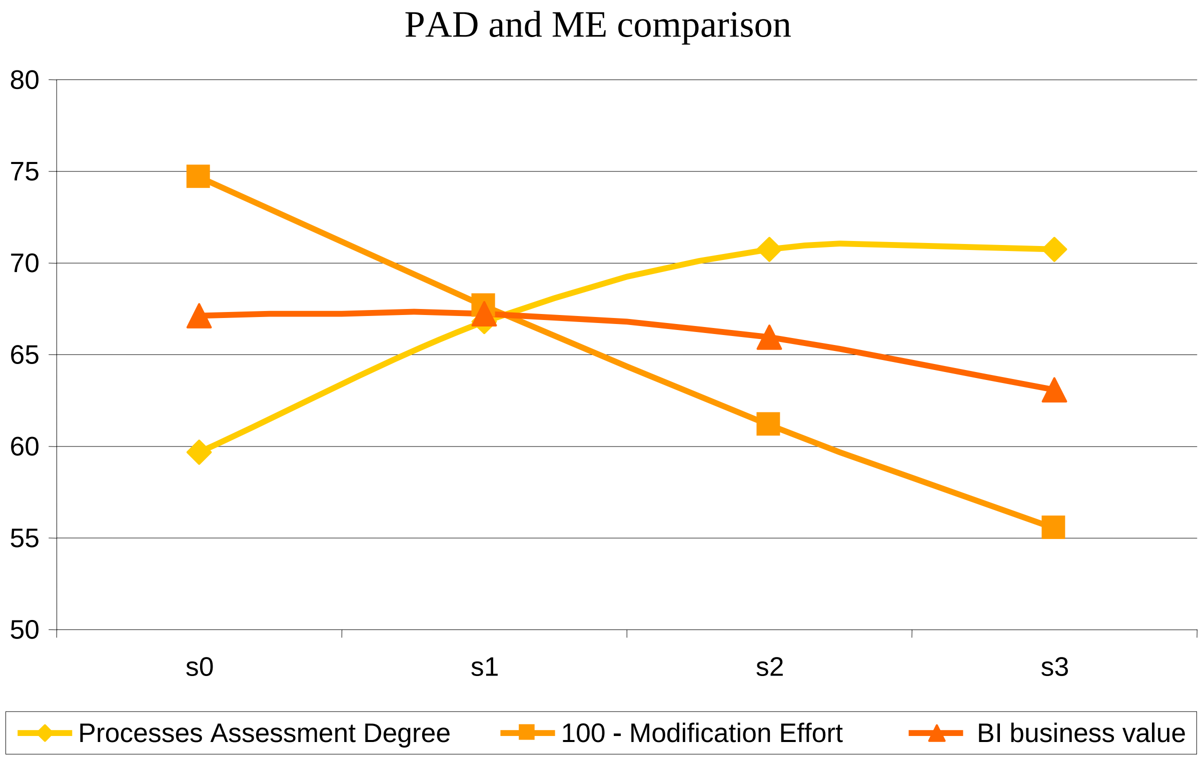


Figure 35, comparison between the values of Processes Assessment Degree and Modification Effort. The Modification effort values have been swapped since it has a negative effect on BI business value. It is possible to see that the resulting BI business value line forms a curve with a differential coefficient close to zero even though it is the sum from two curves with clearly a positive and a negative differential coefficient.

From the diagram above is it possible to see that even though both PAD and ME have two clear tendencies with best and worst scenarios, the result does not have the same clear tendency. Since the two curves of PAD and ME cross each other, and one has a positive differential coefficient, while the other has a negative, they will work to compensate for each other. This is not surprising, since it is the general idea of a cost/benefit analysis, but in this case they compensate each other until all scenarios appear equal and no recommendation can be made.

A second aspect is that the PAD makes such a small difference between the scenarios. The reason is that there are no big differences between how many KPIs that would be made available in different future scenarios. Other changes than those described in scenarios would be needed to increase the PAD too.

The third aspect that will be addressed is the total equality between the scenario 2 and 3 for the PAD. This is of course a very interesting finding since it shows that one would not benefit a bit from choosing scenario 3 instead of 2. But why is it so? This depends mainly on the SNPM which does not recognise any KPIs that use data from outside the nuclear plant and its systems. A Data warehouse has many other benefits but these do not fall inside the delimitations of this study and s3 shows only the same benefits as s2.

A very important thing that should be stressed is that the result can only be compared with itself or with studies done in exactly the same way. A scenario comparison is the ideal way of using this analysing framework but its usefulness towards other areas is limited.

8.2 The nuclear power plant view

This study has completely focused on the cost and the benefits from the nuclear power plant perspective. When it comes to the utility company's costs in general, the costs associated with implementing a DW in the utility company are, for example, not included, except for the efforts that would have to be done by the nuclear power plant. The same applies to the benefit part of the framework. Only the value that would be experienced by the nuclear power plant is part of this study.

This is a very important aspect to remember, especially when thinking about the scenarios, because these issues really concern the whole enterprise. Both the costs and the benefits that would come with scenario 3 are to a great extent not included in this study since these costs and benefits will lie outside of the nuclear power plant.

8.3 Unassessed aspects

The relatively low values in the PAD part of the framework, even for the best PAD scenarios, shows that a lot could be done for the Process Assessment Degree that is not included in the scenarios. During the first interview it also became clear that a lot of the KPIs were not to become available just by implementing scenario 2 or 3. In order to assess a lot of the processes and sub-processes in the SNPM, the whole way of working must be changed. Even though the SNPM is an international standard is it very much influenced by the American ways of working. Therefore, many of the processes in the standard are not directly applicable at Swedish nuclear power plants. A globalisation in the field of nuclear power industry will in time mean that everybody has to adjust a bit in order to fit with international standards.

As a conclusion of this; one of the best ways of increasing the PAD, from the perspective of this study, might be to change the ways of working and the way processes are defined.

8.4 Other attributes of BI systems to consider

As mentioned in *8.2,* the aspects included in this study are only from the nuclear power plant perspective. Furthermore the costs are only analysed from an initialisation perspective. In this chapter, attributes of BI systems that should be considered before taking decisions about future scenarios, which have not been included in the study, will be treated.

8.4.1 Cost attributes

When considering a BI system solution many things except of the modification costs should be considered. In the case when choosing whether to investing in a DW or to settle with local DMs, one need to know that the initial cost for the DW is very high while its maintenance costs are fairly low. The DM has the opposite cost development. [25]

Sofia Sandgren is writing her master's thesis on *Cost assessment for IT scenarios* right now as a project between the utility company and KTH. In her work she recognises Procurement cost, Implementation costs, Operation and Maintenance cost and Termination costs, for future IT scenarios. In her framework she stresses cost of technology, cost of personnel, development costs as well as management cost, through out the lifespan of a system scenario. [46]

This means that depending on if one looks at the investment on a shorter or a longer term, different scenarios could be differently suitable. More research than what has been done in this study is therefore probably needed, maybe covering all costs of each scenario, and maybe with the help of a framework like the one developed by Sandgren.

8.4.2 Beneficial attributes

To only assess the PAD is, even if important, just a very small part of the benefits that come with different choices of system scenarios. Attributes like security and availability are examples of what might be of importance when choosing between scenarios.

Another example of an important attribute is each scenario's ability to have systems fit local needs or organisational standards. Hopefully the whole organisation could standardise their systems towards one way of working, which makes changes/updates so much easier, while still fitting good to the local needs. A good thing is if the system is flexible and can be easily adjusted so that it fits all needs even if the users and the tasks are many.

An attribute of the scenarios very closely related to this study is how different system scenarios enable the information to be accessed. The KPIs in this study for example are always evaluated the same no matter who is given access to them. While in fact the whole point with KPI and performance assessment is to be able to do benchmarks, especially towards others. If the KPI are spread to others which in return spread their KPIs, the KPI value is suddenly much higher. In this case they would have value to so many more. Looking at the scenarios of this study, it is found that only scenario 3 enables a solution which truly makes the KPIs shared by the whole organisation.

To consider other attributes then the ones included in the study it is not only necessary to make a better decision, but it also gives a perspective on the results of this study. Saying for example that the KPIs are going into a whole-organisation-covering DW this would mean the double value for the utility company compared to if the KPIs are staying in the systems at the nuclear power plant. This would mean a drastic change to the value diagram given in *Figure 24*. Note that in the diagram below, what can be seen is not a result of the study, but rather an interesting manipulation of one of the results.

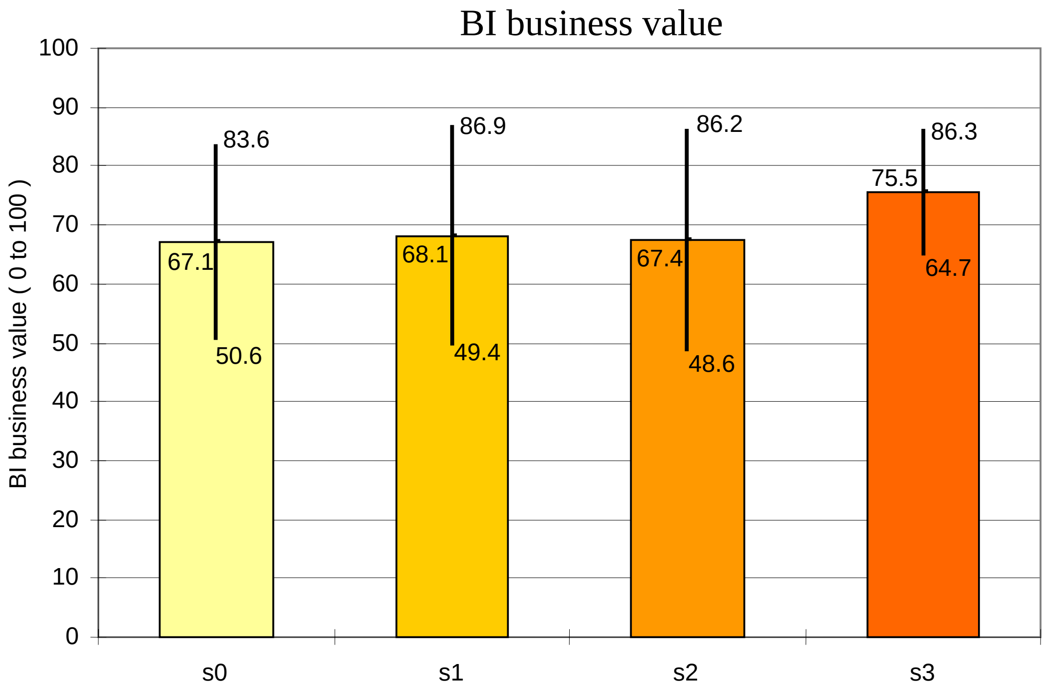


Figure 36, manipulated value diagram of BI business value. The KPIs have been given extra influence if they are available in scenario 3. Note that this is not a result of the actual study.

In the manipulated example shown in the diagram above, Scenario 3 stands out as the best option. The example is done in order to show that other attributes of system solutions can have a great effect on the value of the scenarios and therefore also on the choice of scenario.

8.5 Uncertainty analysis

One of the biggest influences on the result is the uncertainty. In this chapter the sources and the necessity of this uncertainty will be discussed.

The uncertainty grows when there is a part of the framework which is left unanswered or when an answer is not found 100 % credible. This is necessary because it makes the whole study credible when it is possible to see exactly how sure one can be about the result.

There are three main contributors to the uncertainty of the analysis

- Low credibility of empiric data. This means that an answer given during an interview has been assessed to have a low credibility. This can cause the node to get an uncertainty of up to 50 %.
- Unanswered questions. These are nodes which could have been measured but which were not because the right empiric source was not found. This generates a 100 % uncertainty to the affected node in the framework.
- Incomplete framework parts. This means that parts of the framework do not have a measurable end node. It is not possible to know which state is assigned to that part. This generates a 100 % uncertainty for affected nodes. This is in practice only seen in SNPM, which is not a fully completed standard, as described earlier.

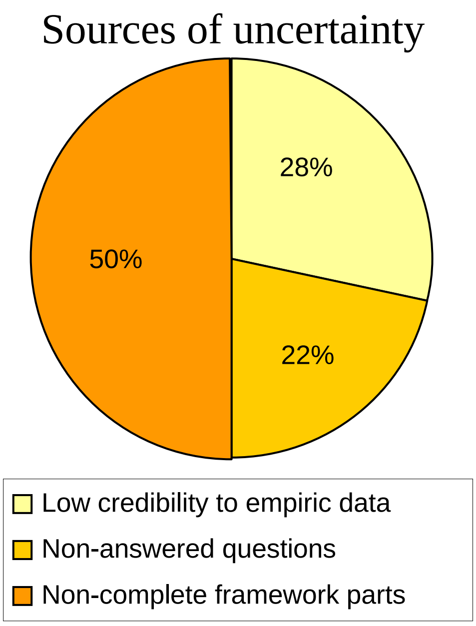


Figure 37, sources of uncertainty of the result. The numbers in this diagram have been calculated from scenario 3, which has an uncertainty 38 % of.

In order to lower the uncertainty and higher the preciseness of the answer one or more of these three parts could be improved.

- Solutions to the first of the three sources of the uncertainty: Persons with better knowledge can be asked the questions with low credibility in a complementary interview, or answers can be confirmed by documents.
- Solutions to the second of the three sources of the uncertainty: The empiric sources which are able to answer the questions can be found. It is also possible that the questions are badly asked and that rephrasing could make them more answerable.
- Solutions to the third of the three sources of the uncertainty, where clearly the highest improvement potential exists: Wait until the next (and hopefully more complete) revision of the SNPM comes, and redo the missing parts. Alternatively find a complementary standard that could replace parts of the SNPM.

A lowering of the uncertainty caused by the incomplete framework parts would also have a direct lowering impact on the unbalanced uncertainty problem described in *7.2 BI business value and Choice of scenario.*

8.6 EID framework

This study has been a pioneer in trying out the EID framework. How well has this framework structure fitted the purpose of the analysis? Well, generally it has worked out according to the expectations. The mathematics behind it is easy to grasp and the straight-forwardness of the EIDs is really one of its greatest advantages. Two limitations in the EIDs, or with GeNIe, which have made cause for some special solutions for this study are:

1. The EID is made to add values through up the structure. The parent always only influences its children. During the work, a general basic idea when constructing the framework was to multiply the inverted value of the Change complexity with the Modifiability, and in this way let Change complexity have the greater influence, the lower

the Modifiability is. Since the EID only allows one to add the values of the parents to the child and not that the parents depend on each other when setting its influences on the child, a special solution with the Modifiability affecting the weight between Modifiability and Change complexity had to be used according to what has been described in *6.2 The framework connections*.

2. A great limitation is that one cannot weigh different scenarios differently. That is, if one would like a different dependency table for different scenarios, this is not possible. During the analysis of the result, it was often tempting to try out what would happen if one applied some special rules for the different scenarios like, the one shown in *Figure 36*. This is however much easier said than done and the values shown in *Figure 36* are a bit of an approximation towards a parallel dependency table for Scenario 3, performed by changing definition values in a simplified copy of the framework.

8.7 Empiric reliability

The empiric work is one of the most essential parts of this study and it is worthwhile to think about whether this important phase has been executed in a good way and what could have been done better.

8.7.1 Connection between the empirics and the analysis

When doing the analysis the answers given during interviews have long since been transformed into numbers, and these numbers have been summed together and multiplied many times. It is very easy to forget where these numbers really came from. What does a 0 actually mean compared to a 50 or a 100.

Especially when completely different parts of a framework are summed together it is important to keep a foot back in reality. Is a 0 the same thing in one part of the framework as in the other? To review this and to scrutinise the connections between the empiric work and the analysis, one should take a close look at *Appendix A* and *Appendix B,* where not only the actual outcome of an answer is to be found, for every end node, but also the criteria from which the answer has been analysed.

8.7.2 The three-graded scale

The three-graded scale that all answers have been quantified to is a key improvement potential for this study. The three grades have proven to be too coarse for the many nuances shown during the interview. Even though the purpose with a simplified framework was good, the actual result has been blindness for the small differences that exist between certain aspects of different scenarios, and an unwanted interviewer-applied control on the answers.

Since the three possible states of each node most cover all possible answers, the most probable answers might only be covered by one of the states, pushing the node to a certain state even before the answer is known. Sometimes, there has in practice only been one possible state, even though the answer could have varied quite some. With a finer granularity of possible states these differences in the answer could have shown.

The number three has proven especially unsatisfying since there are also three different future scenarios. During the Change complexity related interviews, the interviewee was therefore many times tempted to use the coarse scale to show the differences between the scenarios by assigning each of the future scenarios to each of the possible states. This left a feeling that they had not been taking in mind what the states actually meant, and if the answer they gave was really correct by itself, and not just in comparison to the other answers. This is shown in *Figure 33*, where, for example, the second state *Medium* only has clear representation in the second future scenario s2.

8.7.3 Limited amount of sources

The amount of empiric sources was quite low. This is/was due to the relatively tough job of finding the right people to participate in the study. To a limited extent, this meant that persons that are very well qualified to answer some of the questions were asked to answer even some of the questions which they are not as well qualified for answering. This led to a lower credibility of the answers and a higher uncertainty in the result.

8.8 Improvement potential

One of the most important things to do in the end of the study, except of presenting the results, is to identify the areas of improvement potential. This will be done by identifying non-top-ranked nodes for the PAD. For the modifiability, a paragraph identifying the few problems seen generally in the three studied systems will start this chapter. The Change complexity has not been included since this part of the framework only considers future scenarios.

Generally the modifiability in the three studied BI systems is good, but there are a few areas of possible improvement potential. The systems suffer the risk of, what in the framework is called *Decay,* which means that the systems contain old components, and so many changes have already been done on them, that it could imply difficulty to future modifications. All systems have a relatively high level of coupling between different components, which could also lower the modifiability of the systems. The systems are generally well documented everywhere except in the source code it self, where the degree of commenting and the self descriptiveness is very low, making it much harder for other people than the code writer him-/herself to perform modifications. Generally, there is very little help given for the one attempting to do changes. No helping software exists, and the systems are normally not built so that small changes can be implemented more easily.

The following processes could be better assessed by KPIs, and constitutes an improvement potential part of the framework.

Node notification	Processes name	Weight	Value 0 – 1
P2	Equipment Reliability	0.13426	0.5
P1	Manage Configuration	0.11202	0.33333
P5	Nuclear Fuel	0.10157	0.5

Table 18, Processes with improvement potential in the field of performance assessment. Processes are sorted by weight in the framework, so that an improvement in the first of the listed processes would have the most overall improvement effect. The *value 0 – 1* indicates how well the processes are assessed today. The closer the value is to 0 the more it is in need of improvement.

The following Sub-processes could be better assessed by KPIs and constitute an improvement potential part of the framework

Node notification	Processes name	Weight	Value 0 – 1
SP33	WM004 – Perform Corrective Maintenance	0,03665	0
SP2	CM002 – Change Design Requirements	0,0365	0
SP26	OP003 – Monitor and Control Plant Chemistry	0,03633	0,75
SP24	OP001 – Operate and Monitor Structures, Systems and Components	0,03309	0,2
SP14	LP006 – Provide Fire Protection	0,02996	0,25
SP16	MS002 – Procure Materials	0,02902	0,375
SP3	CM003 – Change Physical Configuration	0,02749	0,6
SP11	LP003 – Provide Safety Services	0,02659	0,5
SP31	WM002 – Perform Scheduling	0,02509	0
SP32	WM003 – Perform Preventive Maintenance	0,02472	0,5
SP30	WM001 – Perform Planning	0,02325	0
SP38	WM009 – Perform Minor Maintenance/ Fix-It-Now Maintenance	0,02227	0
SP1	CM001 – Evaluate Identified Problem or Desired Change	0,0211	0,125
SP8	ER004 – Perform Predictive Maintenance	0,02053	0,5
SP18	MS004 – Provide Warehousing	0,01767	0,6875
SP4	CM004 – Change Facility Configuration Information	0,01713	0
SP27	T001 – Develop Training Programs	0,01467	0

Table 19, Sub-processes with improvement potential in the field of performance assessment. Sub-processes are sorted by weight in the framework, so that an improvement in the first of the listed sub-processes would have the most overall improvement effect. The *value 0 – 1* indicates how well the processes are assessed today. The closer the value is to 0 the more it is in need of improvement.

An interesting thing to study is how well the Processes and Sub-process which by the nuclear power plant have been given a higher weight are assessed. Looking at the processes, it turns out that KPIs relating to the highest weighted processes are normally very well assessed. When it comes to the Sub-process, though, the majority of the well assessed KPIs are related to low weighted Sub-processes. For the processes, this has a positive affect on the BI business value, while for the Sub-processes it will affect the BI business value negatively.

9 Future work

Even though this study is now completed, there are many potential follow-up works that could be interesting to pursue in the future.

The spread of uncertainty of the study is relatively high, and, as discussed in *8.5 Uncertainty analysis,* there are many different ways to narrow down this spread. A future study could focus on asking questions about the uncovered parts of the SNPM. Or it could ask questions about the Modifiability of SAP R/3 to a person with better knowledge. During the study Monica Sörerbom was sought after, without success, for the purpose of strengthening this very low credibility part of the study, and in a future work she might be an interesting source.

A future work could also choose to study the many non-assessed aspects of BI. It would be especially interesting to look upon the questions of this study more from an enterprise perspective, and to not just look at which benefits and costs that the nuclear power plant would enjoy/suffer from.

It would also be interesting to look upon benchmarking and information sharing from a more strategic perspective. Which are the long time affects of having a unified way of measuring processes and benchmarking within and outside of the own organisation? What is the company strategy in this matter? How do other players on the energy market act in these questions?

The cost aspects of future scenarios could also be greatly further studied. In this study, only the initial implanting costs have been assessed. An interesting follow-up study could use the framework for assessing the total cost of IT scenarios provided by Sandgren [46] and measure the scenarios used in this study to get a more complete picture.

10 Acknowledgements

The following persons should be acknowledged for their important contributions to the completion of this master thesis. 13 persons have preferred to be anonymous and are not presented here even though the success of study has been dependent on their contribution.

- Pontus Johnson, KTH, Examinator
- Robert Lagerström, KTH, Mentor
- Daniel Oscarsson, KTH
- Sofia Sandgren, KTH
- Jonas Söderbom, KTH
- Joseph Schippert, MRO software
- Charlotte Uggla, Stockholms stad
- Johan Ullberg, KTH
- Jonas Öhrström, KTH

11 Bibliography

[1] Arnoldsson, A., *Förslag till principbeslut Datawarehous för the utility companykoncernen*. 2003, the utility company. p. 1-5.

[2] Banker, R.D., et al., *Software complexity and maintenance costs.* Commun. ACM, 1993. 36(11): p. 81-94.

[3] Bengtsson, P., et al., *Architecture-level modifiability analysis (ALMA).* J. Syst. Softw., 2004. 69(1-2): p. 129-147.

[4] Bezzina, D., *Data quality in Business Intelligence Systems Presentation*, G. Lindmark, Editor. 2006, SAS Institute: Stockholm. p. Company presentation.

[5] Bose, R., *Understanding management data systems for enterprise performance management* Understanding management data systems for enterprise performance management 2006. 106(1): p. 43-59.

[6] Chan, A.P.C., *Key performance indicators for measuring construction success.* Benchmarking: An International Journal, 2004. 11(2): p. 203-221.

[7] Chan, T., S.-L. Chung, and T.H. Ho, *An Economic Model to Estimate Software Rewriting and Replacement Times.* IEEE Trans. Softw. Eng., 1996. 22(8): p. 580-598.

[8] Chaudhuri, S., *An Overview of Data Warehousing and OLAP Technology.* ACM Sigmod Record, 1997.

[9] Cody, W.F., *The integration of business intelligence and knowledge management.* IBM Systems Journal, 2002. 41(4): p. 697-713.

[10] DSL, *GeNIe 2.0*. 1998, Decision Systems Labratory at the University of Pittsburg: Pittsburg. p. Data program.

[11] Eriksson, M. and J. Lilliesköld, *Projekthandbok för mindre projekt*. 2004, Stockholm: Liber.

[12] Feinberg, D. and M.A. Beyer, *Magic Quadrant for Data Warehouse Database Management Systems*, Gartner, Editor. 2006, Gartner RAS core Research Note.

[13] the nuclear power plant, *En detaljerad beskrivning av områdena affärsplanering och uppföljning i affärsprocessen*. 2006.

[14] the nuclear power plant, *Manual till Aktuell prognos och anslutande filstruktur*. 2006.

[15] Gammelgård, M., *Utredningsmetodik I, Industriella Informations och styrsystem, KTH*, G. Lindmark, Editor. 2006, Department of Industrial Information and Control Sytems Royal Insitute of Technology (KTH): Stockholm. p. Presentation in case study methology.

[16] Gammelgård, M., *Utredningsmetodik II, Industriella Informations och styrsystem, KTH*, G. lindmark, Editor. 2006, Department of Industrial Information and Control Sytems Royal Insitute of Technology (KTH): Stockholm. p. Presentation in case study methology.

[17] Gardner, S.R., *Building the Data Warehouse.* Communications of the ACM, 1998. 41(9): p. 52-60.

[18] Golfarelli, M., S. Rizzi, and I. Cella, *Beyond data warehousing: what's next in business intelligence?*, in *Proceedings of the 7th ACM international workshop on Data warehousing and OLAP*. 2004, ACM Press: Washington, DC, USA.

[19] Graves, T.L. and A. Mockus, *Inferring Change Effort from Configuration Management Databases*, in *Proceedings of the 5th International Symposium on Software Metrics*. 1998, IEEE Computer Society.

[20] Hale, J., et al., *Enhancing the Cocomo Estimation Models.* IEEE Softw., 2000. 17(6): p. 45-49.

[21] ICS. *Forsknings områden*. 2006 [cited 2006-11-25]; Available from: http://www.ics.kth.se/Forskning/forskning_områden.htm.
[22] Inmon, W.H., *Building the Data Warehouse*. Third Edition ed. 2002: John Wiley & Sonns, Inc. 412.
[23] ISO, S., *ISO 9126*. 2001. 1-4.
[24] Johnson, P., et al. *Extended Influence Diagrams for Enterprise Architecture Analysis*. in *Proceedings of the 10th IEEE International Annual Enterprise Distributed Object Computing Conference*. 2006: Department of Industrial Information and Control Sytems, Royal Insitute of Technology.
[25] Jonasson, D., *Data Warehouse and Data Mart - Användning, tillämpbarhet och rekomendationer*, in *School of Mathematics and System Engineering*. 2004, Växjö University: Växjö. p. 49.
[26] Jones, A.E., *Data Warehouses, Data Marts, OLAP, Data Mining*. 2005, Staffordshire University, School of Computing. p. 20.
[27] Karlén, G., *Conversation with G K and J M*, G. Lindmark, Editor. 2006: Stockholm.
[28] Kudyba, S., *Data mining and business intelligence : a guide to productivity*. 2001, Hersey: Idea Group Publishing. 166.
[29] Lagerström, R., *Assessing Modifiability - A method applied at the utility company AB*. 2005, Industrial Information and Control Systems, Royal Institute of Technology
[30] Lagerström, R., *EID Consolidation Process*. 2006, Department of Industrial Information and Control Sytems Royal Insitute of Technology (KTH).
[31] Lagerström, R. and P. Johnson, *Extended Influence Diagram Generation for Interoperability Analysis, To Appear*. 2006, Department of Industrial Information and Control Sytems Royal Insitute of Technology.
[32] Luhn, H.P., *A Business Intelligence System*. IBM Journal, 1958: p. 314 - 319.
[33] Mathworks, *Matlab*. 2005, Mathworks Inc.
[34] Merriam, S.B., *Fallstudien som forskningsmetod*. 1994: Studentlitteratur AB. 228.
[35] Microsoft, *Microsoft Office Excell*. 2003, Microsoft Corporation.
[36] Miyazaki, Y. and K. Mori, *COCOMO evaluation and tailoring*, in *Proceedings of the 8th international conference on Software engineering*. 1985, IEEE Computer Society Press: London, England.
[37] Mockus, A. and L.G. Votta, *Identifying Reasons for Software Changes Using Historic Databases*, in *Proceedings of the International Conference on Software Maintenance (ICSM'00)*. 2000, IEEE Computer Society.
[38] Mützell, J., *Conversation with J M*, G. Lindmark, Editor. 2006: Stockholm.
[39] NEI, *The Standard Nuclear Performance Model - A, Process Management Approach - Revision 4*, N.E.I.N.A.M.C.o.P. Report, Editor. 2004, Nuclear Energy Institute: Washington D. C. p. 29.
[40] NEI. *Nuclear Energy Institute*. 2006 [cited 2006-11-01]; Available from: http://www.nei.org/.
[41] Nemati, H.R., *Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing*. Decision Support Systems, 2002. 33: p. 143 - 161.
[42] Oman, P., J. Hagemeister, and D. Ach, *A Definition and Taxonomy for Software Maintainability*. Software Engineering Lab, Hewlett-Packard, 1992.
[43] Persson, R., *Business Intelligence - the customers view*, in *Division of Production Management*. 2006, Lund University: Lund. p. 1-76.
[44] Reh, F.J. *Key Performance Indicators*. 2006 2006-10-23 [cited; Available from: http://management.about.com/cs/generalmanagement/a/keyperfindic.htm.

[45] Roast, C.R. and J.I. Siddiqi, *Relating Knock-on Viscosity to Software Modifiability*, in *Proceedings of the 6th Australian Conference on Computer-Human Interaction (OZCHI '96)*. 1996, IEEE Computer Society.
[46] Sandgren, S., *Cost assesment for IT scenarios*. 2006, Department of Industrial Information and Control Sytems Royal Insitute of Technology (KTH). p. 48.
[47] Schippert, J., *Mail: RE: Which KPI's does MRO recognize in Nuclear Power Generation*, G. Lindmark, Editor. 2006, MRO software. p. Mail conversation.
[48] SKI. *Statens Kärnkraft Inspektion*. 2006 [cited 2006-11-01]; Available from: http://www.ski.se/.
[49] Sörqvist, L., *Ständiga Förbättringar*. 2004: Studentlitteratur. 572.
[50] Thierauf, R.J., *Effective Business Intelligence Systems*. 2001, Westport: Quorum Books.
[51] WANO, *WANO 2005 Perfomance Indicators*. 2005, World Association of Nuclear Operations: London.
[52] WANO. *What is WANO?* 2006 [cited 2006-11-01]; Available from: http://www.wano.org.uk/WANO_Documents/What_is_Wano.asp.
[53] Wheeler, A. and L. Wheeler. *Knowledge base technology*. 2006 [cited 2006-11-25]; Available from: http://www.puffin.com/~lynn/secgloss.htm#t181.
[54] Yin, R.K., *Case Study Research: Design and Methods Third Edition* 2003: Sage Publications, Inc. 200.
[55] Ö, B.-E., *Feasability Report: Data Warehouse For MIS/DSS* 2003, the utility company Data: Stockholm. p. 1-39.

12 Appendixes

Appendix A	Description of the ME framework including collected empiric data
Appendix B	Description of the PAD framework including collected empiric data
Appendix C	Creation of the EIDs
Appendix D	Further concepts of Business Intelligence

Please note that the references given in the appendixes refer to the bibliography given in the end of each appendix document.