Evaluation of Knowledge Management Technologies for Strategic Innovation Management

Prof. Dr. Ludwig Nastansky

Wintersemester 2004/2005

vorgelegt von:

Marc Henselewski
Wirtschaftsinformatik
Matr.-Nr. 6120620
Gräffstrasse 32
45894 Gelsenkirchen

in Kooperation mit

DETECON Consulting
Danksagung:

Allen, die mir mit Rat und Tat bei meiner Diplomarbeit behilflich waren, möchte ich an dieser Stelle meinen herzlichsten Dank aussprechen.


Aber auch ohne die freundliche Unterstützung der vielen hier nicht namentlich aufgeführten Personen, die mir Auskunft gaben, Rat erteilten, Korrektur lasen oder in sonstiger Weise zu dem Gelingen dieser Arbeit beigetragen haben, hätte diese Arbeit ungleich langsamer und schwieriger Gestalt angenommen.
Eidesstattliche Erklärung:

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Paderborn, den .......... ................
(Datum) (Unterschrift)
# Table of Contents

4 EVALUATION OF KNOWLEDGE MANAGEMENT TECHNOLOGIES ............................................. 35
   4.1 DEVELOPMENT OF A CHARACTERIZATION SCHEME FOR THE EVALUATION ............................................. 36
      4.1.1 The Technology Forecasting Dimension ................................................................. 36
      4.1.2 The Knowledge Management Dimension ............................................................. 38
      4.1.3 The Context-Complexity-Matrix ............................................................................. 41
   4.2 EVALUATION ......................................................................................................................... 42
      4.2.1 Data Mining .................................................................................................................. 43
         4.2.1.1 Introduction and Definition of Important Terms ....................................................... 43
         4.2.1.2 Methods and Application Scenarios ........................................................................ 45
         4.2.1.3 Analysis and Characterization .................................................................................. 50
      4.2.2 Case-Based Reasoning .................................................................................................... 51
         4.2.2.1 Introduction and Definition of Important Terms ....................................................... 51
         4.2.2.2 Methods and Application Scenarios ........................................................................ 55
         4.2.2.3 Analysis and Characterization .................................................................................. 57
      4.2.3 Information Retrieval ..................................................................................................... 59
         4.2.3.1 Introduction and Definition of Important Terms ....................................................... 59
         4.2.3.2 Methods and Application Scenarios ........................................................................ 61
         4.2.3.3 Analysis and Characterization .................................................................................. 65
      4.2.4 Topic Maps ...................................................................................................................... 67
         4.2.4.1 Introduction and Definition of Important Terms ....................................................... 67
         4.2.4.2 Methods and Application Scenarios ........................................................................ 71
         4.2.4.3 Analysis and Characterization .................................................................................. 73
      4.2.5 Ontologies ...................................................................................................................... 75
         4.2.5.1 Introduction and Definition of Important Terms ....................................................... 75
         4.2.5.2 Methods and Application Scenarios ........................................................................ 77
         4.2.5.3 Analysis and Characterization .................................................................................. 80
   4.3 INTEGRATIVE DISCUSSION OF THE EVALUATION ............................................................. 83

5 A REAL WORLD EXAMPLE ........................................................................................................ 86
   5.1 EXAMPLE SCENARIO: DETECON, INC. – DEUTSCHE TELEKOM AG ................................. 86
      5.1.1 Business Process Description ....................................................................................... 86
      5.1.2 Process Evaluation ........................................................................................................ 88
   5.2 SOLUTION CONCEPT .......................................................................................................... 89
      5.2.1 Organizational Issues ..................................................................................................... 90
      5.2.2 Knowledge Management System Concept ...................................................................... 91
   5.3 DISCUSSION ......................................................................................................................... 93

6 CONCLUSION ............................................................................................................................. 94
   6.1 OUTLOOK ............................................................................................................................. 94
   6.2 SUMMARY ............................................................................................................................ 95

BIBLIOGRAPHY .............................................................................................................................. 97
Figures

FIGURE 2-1 STAGES OF FORECASTING ........................................................................................................ 13
FIGURE 2-2 TECHNOLOGY FORECASTING PROCESS FOR STRATEGIC INNOVATION MANAGEMENT ..... 15
FIGURE 2-3 METHODOLOGY TREE BY ARMSTRONG (DOTTED LINES PRESENT POSSIBLE RELATIONSHIPS) .. 16
FIGURE 2-4 THE SECI PROCESS (ADAPTED FROM NONAKA, TAKEUCHI 1995) ................................. 24
FIGURE 3-1 ROLES WITHIN TECHNOLOGY FORECASTING (ADAPTED FROM MILES, KEENAN, KAIKO-OJA 2002, P. 163) ........................................................................................................ 27
FIGURE 4-1 INFORMATION-STRUCTURE COMPLEXITY OVER THE TECHNOLOGY FORECASTING PROCESS... 37
FIGURE 4-2 THE DATA-INFORMATION-KNOWLEDGE MODEL (AAMODT, NYGÅRD 1995, P. 8) .......... 38
FIGURE 4-3 THE CONTINUUM OF CONTEXT EXPLICATION (SMOLNIK, KREMER, KOLBE 2005, P. 38) .... 40
FIGURE 4-4 THE CONTEXT-COMPLEXITY MATRIX ...................................................................................... 42
FIGURE 4-5 THE KDD PROCESS (ESTER, SANDER 2000, P. 2) .................................................................. 45
FIGURE 4-6 EXAMPLE REPRESENTATIONS OF A CLASSIFIER (ESTER, SANDER 2000, P. 108) ............... 48
FIGURE 4-7 THE CBR-CYCLE (AAMODT, PLAZA 1994, P. 45) ................................................................. 53
FIGURE 4-8 THE MODEL OF INFORMATION RETRIEVAL (INGWersen 1992, P. 55) ............................... 61
FIGURE 4-9 TOPICS AND OCCURRENCE ROLES (PEPPER 2002, P. 9) .................................................. 69
FIGURE 4-10 TOPICS, OCCURRENCE ROLES, AND ASSOCIATIONS (PEPPER 2002, P. 10) ................. 70
FIGURE 4-11 WEB PAGE ANNOTATION FOR ONTOLOGY REFERENCES ............................................. 80
FIGURE 4-12 COMPLETED CONTEXT-COMPLEXITY MATRIX ............................................................... 84
FIGURE 4-13 REVISED FOCUS WITHIN THE CONTEXT-COMPLEXITY MATRIX ............................... 85
FIGURE 5-1 TECHNOLOGY FORECASTING PROCESS AT DETECON, INC. FOR DEUTSCHE TELEKOM AG.... 88
FIGURE 5-2 KNOWLEDGE MANAGEMENT SYSTEM CONCEPT FOR DETECON, INC. ...................... 93
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>API</td>
<td>Application Programming Interface</td>
</tr>
<tr>
<td>CBR</td>
<td>Case-based Reasoning</td>
</tr>
<tr>
<td>DAML</td>
<td>DARPA Agent Markup Language</td>
</tr>
<tr>
<td>DARPA</td>
<td>Defense Advanced Research Projects Agency</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>IR</td>
<td>Information Retrieval</td>
</tr>
<tr>
<td>KDD</td>
<td>Knowledge Discovery in Databases</td>
</tr>
<tr>
<td>MOP</td>
<td>Memory Organization Package</td>
</tr>
<tr>
<td>OIL</td>
<td>Ontology Inference Layer</td>
</tr>
<tr>
<td>OLAP</td>
<td>Online Analytical Processing</td>
</tr>
<tr>
<td>R&amp;D</td>
<td>Research and Development</td>
</tr>
<tr>
<td>RDF</td>
<td>Resource Description Framework</td>
</tr>
<tr>
<td>SGML</td>
<td>Standard Generalized Markup Language</td>
</tr>
<tr>
<td>SQL</td>
<td>Structured Query Language</td>
</tr>
<tr>
<td>W3C</td>
<td>World Wide Web Consortium</td>
</tr>
<tr>
<td>WWW</td>
<td>World Wide Web</td>
</tr>
<tr>
<td>XML</td>
<td>Extended Markup Language</td>
</tr>
<tr>
<td>XTM</td>
<td>XML Topic Maps</td>
</tr>
</tbody>
</table>
1 Introduction

1.1 Scenario

The competition in today’s business environment is growing faster than ever before. Influences of the high pace with which globalization moves forward, and the ongoing liberalization of national and international markets lead to the emergence of new problem settings and increasing pressure for existing companies. The extension of the European Union and the creation of an open European market is just one example for this development, which creates new business opportunities on the one hand, but also augments the number of competitors in the market on the other. Thus, companies are facing greater risks due to a higher number of players in the market. Important business decisions have to be made in a shorter time frame and with a wider spectrum of available information in order to achieve or sustain competitive advantage.

However, environmental influences which are created externally of the market are not the only factors that have an impact on the complexity of the situation of existing companies. The increasing speed with which innovations and new developments are being made also adds to the pressure felt by the firms and their decision-makers through shorter product life cycles and decreasing production costs. High technology companies in particular, who have enormous expenditure for research and development (R&D), have to plan their research programs more carefully, because of a higher risk of losing competitive advantage when “going the wrong way”. As a consequence, decision-makers have a greater need to anticipate or forecast future developments and apply these insights to business strategies and strategic innovation management to keep risk levels low and the company in the competition. According to Bright all “…firms and governments dealing with technology have been and are doing technology forecasting. This is because each decision to explore, support, oppose or ignore a technological prospect incorporates the decision-maker's assumptions about that technology and its viability in the future" (Bright 1979, p.228). Moreover, companies have the need to formalize and structure those efforts in their innovation management to improve the quality of the decision-making process and to achieve a technology leadership position.

The aforementioned needs become even more obvious when thinking about multinational enterprises. These companies compete on several markets at the same time and their R&D departments are often distributed over different continents. It is
nearly impossible to control the various research activities without possessing a company-wide and future-oriented R&D strategy. To be able to create such a strategy, decision-makers have to assess a myriad of different information in a very limited time frame.

For the last several years, firms have been realizing more and more that knowledge plays a key role in the development of strategies for future success and stronger market positions. The most striking examples are technology and service-oriented companies, but retailers also engage in activities to use knowledge and information as factors of competitive advantage. A paradigm shift of business strategies from a focus on tangible assets to one which prioritizes intangible assets can be observed. However, the quantity of information and information sources increases continuously and what first seemed to be the solution for several business problems became a unique problem itself for the companies of today; namely, too much information. In order to gain from information and to facilitate knowledge creation within a company, new ways of filtering and selecting information have to be applied.

Furthermore it is obvious that the nature of knowledge is highly dynamic. The value of knowledge is difficult to measure and can change from one moment in time to another. Companies are trying to control this uncertainty to some extent and to get as much advantage out of their companies’ knowledge as possible by integrating knowledge management aspects into competitive strategies.

### 1.2 Problem Setting

As described previously, circumstances of today’s companies can be characterized by high competitive pressure, a tendency towards a prioritization of intangible assets, an overwhelming amount of information and information sources, and a stronger need to anticipate developments and changes in the future. The last point in particular becomes a major aspect of competitive advantage for firms in high-technology industries.

It is the objective of this thesis to suggest a solution concept to support technology forecasting in a high competitive environment to facilitate strategic innovation management and decision making with the help of knowledge management technologies\(^1\). To achieve the objective, the following goals will be addressed:

- Analysis of technology forecasting with respect to knowledge management.

---

\(^1\) sometimes also referred to as Knowledge Management Methods or Instruments
Definition of a characterization scheme for the evaluation of capabilities of knowledge management technologies with respect to supporting technology forecasting and its application to selected technologies.

Introduction of a solution concept for improved strategic innovation management support, described with the help of a real world example.

1.3 Structural Overview of the Thesis

The discussion of the goals starts with a brief overview of the area of strategic innovation management to facilitate understanding of the focus of the thesis and the importance of technology forecasting. This is followed by an individual introduction of technology forecasting and knowledge management. Both fields of interest are defined and examined in detail, in order to provide a comprehensive theoretical base for the subsequent chapters.

Chapter three ties in with chapter two where the definitions and theories presented in the previous chapter are used to analyze technology forecasting with respect to knowledge management. First, the need for knowledge management within technology forecasting will be determined. Therefore, an approach is chosen which differentiates between knowledge management needs emerging from inside the company and those emerging from outside the company. Following that, the process of technology forecasting is analyzed step by step in order to characterize the information and knowledge used and to identify the potential of performance improvement through knowledge management technologies for each single step of the process.

A characterization scheme is developed at the beginning of chapter four to facilitate categorizing of knowledge management technologies. The suggested scheme is used in combination with the earlier developed insights to evaluate selected knowledge management technologies. For each technology the potential of supporting technology forecasting is determined and the optimal way to exploit this potential identified.

After pointing out which of the selected knowledge management technologies might deliver the highest value for specific parts of technology forecasting, a real world example is used in chapter five to apply the insights of the preceding chapters to the creation of a concrete solution concept in the field of strategic innovation management.
The concluding chapter six gives an outlook on further research needs and raises questions which could not be answered within the scope of this thesis. It closes with summarizing the major findings and provides a critical review of the thesis.

2 Delimitation and Conceptual Definitions

This chapter will provide the necessary theoretical base for the subsequent chapters. Each of the three fields which are combined through the focus of the thesis, i.e. strategic innovation management, technology forecasting, and knowledge management, will be introduced, described and important terms defined.

2.1 Strategic Innovation Management

An overview of the most important terms and aspects of strategic innovation management, which plays a major role in a company’s success, will be given in this section. First the term *innovation* will be defined after which strategic innovation management will be addressed. For more detailed information about the field of strategic management or innovation management, the reader is referred to the corresponding literature; see especially Porter’s “Competitive Advantage” or Burgelman, Maidique, and Wheelwright’s “Strategic Management of Technology and Innovation”.

2.1.1 Definition of Innovation

The term *innovation* is commonly used within many situations, for example with respect to technological progress or advertisement. We all have heard the slogan “innovative design” at least once. But in the context of business and economics, the term *innovation* has to be defined more precisely. Allesch, for example, defines innovation as

“*the efficient and effective transformation of new technologies into marketable products*”\(^1\) (Allesch 1986, p. 3)

The definition above contains two important aspects about innovation. Firstly, innovation can be seen as a process and not only as resulting products. Secondly, it is essential for the definition of innovation to emphasize the development of a marketable product. That is, innovation is the exploitation of new technologies in order to achieve a higher gain in terms of sales and revenues. In order to achieve that, the developed product has to create new customer value or higher customer value than preceding

---

\(^1\) As an emphasis all definitions are in italics. Therefore the format of quotes might differ from originals.
products. O’Hare formulates his definition of innovation entirely on the basis of this thought:

“New way of delivering customer value” (O’Hare 1988, p. 27)

In addition, O’Hare extends the discussion about innovation and argues that “truly successful innovation does not just lead to some extra sales volume, or a temporary improvement in performance. ... Rather, it is about achieving fundamental improvement in competitive position, about re-establishing the competitive equilibrium at a new, more favorable point. To do this it must not only create value for customers, but must do so in a way which gives the innovator a defensible competitive advantage.” (O’Hare 1988, pp. 39-40). This is underlined by Geschka, who states that “the innovation strategy is one means to achieve overall strategic company goals” (Geschka 1992, p. 70).

Furthermore, it can be discussed that different types of innovation exist. According to Burgelman, Maidique, and Wheelwright, innovation can be subdivided into three different types:

- “Incremental innovations involve the adoption, refinement, and enhancement of existing products and services and/or production and delivery systems.”
- “Radical innovations involve entirely new products and service categories and/or production and delivery systems.”
- “Architectural innovations refer to reconfigurations of the system of components that constitute the product - for example, the effect of miniaturization of key radio components.” (Burgelman, Maidique, Wheelwright 1996, p. 2)

All in all, the following definition of innovation can be derived from the information presented in this section:

Innovation comprises the adoption, refinement, and enhancement of existing products, services and/or production and delivery systems, the transformation of new technologies, or the reconfiguration of system components that constitute a product, with the aim of developing a new marketable product which creates new ways of delivering customer value and establishes, sustains or improves competitive advantage.
2.1.2 Definition of Strategic Innovation Management

When approaching strategic innovation management, it is useful to start by clarifying the general meaning of the term *management*.

“*Organization, supervision, or direction; the application of skill or care in the manipulation, use, treatment, or control (of a thing or person), or in the conduct of something*” (Oxford English Dictionary, Second Edition 1989)

As the basic definition of management implies, the term management incorporates planning, monitoring, and directing, which is also true in the case of strategic management. “Strategic management, often called ‘policy’ or nowadays simply ‘strategy’, is about the direction of organizations, and most often, business firms.” (Rumelt, Schendel, Teece 1991, p. 5). Pearce and Robinson use a more general approach with respect to the activities included into strategic management and define strategic management as

“the set of decisions and actions that result in the formulation and implementation of plans designed to achieve a company's objectives.”

(Pearce, Robinson 2000, p. 3)

Apparently, all activities within strategic management have the purpose to align a company’s functions and capabilities toward a defined set of goals and objectives. Besides formulation and implementation of plans, David also takes the evaluation of plans and strategies into account and states:

“*Strategic management can be defined as the formulation, implementation, and evaluation of actions that will enable an organization to achieve its objectives.*” (David 1986, p. 4)

David explains that formulation involves identifying an organization’s internal strengths and weaknesses, determining a firm’s external opportunities and threats, establishing a company mission, setting objectives, developing alternative strategies, analyzing these alternatives, and deciding which ones to execute, while implementation comprises establishing goals, devising policies, motivating employees, and allocating resources in a manner that will allow formulated strategies to be pursued successfully. Evaluation, in the end, includes all activities necessary for monitoring the results of strategy formulation and implementation (David 1986, p. 4).
Keeping David’s definition in mind, it is possible to derive a definition for strategic innovation management with the help of the discussion of innovation presented in the preceding section. Strategic innovation management has to be seen as a sub-activity of strategic management, since it takes the formulated strategies as an input in order to plan objectives for innovation and the R&D program. Geschka mentions that “the targets for innovations are derived from such overall goals; they are in accordance and linked to other partial strategies of the company” (Geschka 1992, p. 70). Thus, strategic innovation management can be defined as

\[ \text{the formulation, implementation, and evaluation of strategies and actions that align innovation with an organization’s competitive strategy and overall objectives.} \]

2.2 Technology Forecasting

Technology forecasting forms the focus of this thesis. Within this section an introduction to its motivation, definitions, and processes will be given. Additionally, forecasting will be delineated from the related term **foresight** and an overview of forecasting methods will be presented.

2.2.1 Motivation and History

After the beginning of the industrial revolution in the mid-1700s a great amount of literature about technology and the future appeared for various different reasons. Amongst these pieces of literature were philosophical essays, utopian tracts, or warnings using technical reports and scientific studies as background information which can partially be compared to the forecasts of today (Bright 1979, p. 229). Since the mid-1900s, the number of forecasting related activities increased faster through the emergence of institutions that were and still are “devoted to supporting their societal role by advancing their technology” (Bright 1979, p. 229), e.g. defense, agriculture, medicine, or communications. The first time technology forecasting appeared in an academic context in the USA was in the 1880s.

After the Second World War technology forecasting became an even more important aspect of organizational decision making processes and strategy development. Bright dates the first major organizational effort to forecast technology to facilitate management decisions in the USA to 1945 which was an Army Air Forces-undertaking in order to achieve guidance in planning research and development (Bright, 1979, p.
In the following years the growth of the community of technology forecasters got faster, especially through the 1960s and 70s. “None of these activities was new, but they had ‘… never before been put forward by so vociferous group of people. Looking back this period is easily recognized as the era of ‘grand futurism’” (Loveridge 1997, p.1).

Today, forecasting is used for many different reasons. Every decision with effects on the future can be supported by forecasting principles, like decisions in politics, production plans, or procurement. Other example situations in which forecasting activities are conducted include

- estimation of future sales of a set of products to ensure that production and inventory are kept at economical levels,
- determination of returns on investment, or
- analysis of different scenarios of economic variables, such as unemployment, in order to compare different political strategies (more detailed information on usage scenarios can be found in DeLurgio 1998, p. 7).

History and the examples mentioned above indicate that the main motive for the use of forecasting is uncertainty of future developments, mostly tied to a decision situation. The more uncertain the circumstances of a decision and the higher the risks which come with it are, the greater is the need for a comprehensive view about what might happen in the future. Armstrong says about forecasting and uncertainty that “we have no need to forecast whether the sun will rise tomorrow. There is also no uncertainty when events can be controlled; for example, you do not need to predict the temperature in your home. Many decisions, however, involve uncertainty, and in these cases, formal forecasting procedures … can be useful” (Armstrong 2001, p. 2).

### 2.2.2 Definition of Technology Forecasting

As has been shown above, forecasting is used in many different situations, in all of which a decision-maker has to deal with some amount of uncertainty. Nevertheless, uncertainty is not the only defining element of forecasting. This section explores the different aspects and requirements of forecasting in greater detail and tries to devise a thorough general definition of forecasting. Once this general definition is complete, it will then be rendered more precisely in order to develop a definition of technology
forecasting. Additionally it will be shown how forecasting can be distinguished from foresight, which also plays an important role in the field of futures studies.

2.2.2.1 Forecasting

Since forecasting itself is a very common term and often used in everyday life, e.g. weather forecasts, all of us have some idea about what forecasting stands for. This common use, however, does not lead to a simple and unique definition of the term. Armstrong, for example, argues that “forecasting is often confused with planning. Planning concerns what the world *should* look like, while forecasting is about what it *will* look like” (Armstrong 2001, p. 2), just to mention one example for possible confusion related to forecasting.

Up to this point the reader should be aware of the fact that forecasting deals with reducing uncertainty for decisions about future developments. But when thinking about future developments it is necessary to clearly specify the determining factors of the future which are considered; in other words, the timing or horizon length of a forecast has to be determined before any forecasting activities can be conducted. This is also important for the actual degree of uncertainty with which a forecaster is confronted. Reconsidering the mentioned example of weather forecasts, it is obvious that the chances for forecasts of tomorrow’s weather to be correct are a lot higher than for forecasts of the weather conditions in two or three weeks. With an increasing horizon length the accuracy of the results of the forecasting activities decreases. Another example for the relation between the accuracy of a forecast and its horizon length is material-demand forecasting in production processes (DeLurgio 1998, p. 9).

To be able to deal with uncertainty during a planning or decision situation, decision-makers need to be prepared for a variety of contingencies and scenarios. It is clear that not all of the possible contingencies and scenarios have the same probability to actually become a reality and hence, priorities for creating plans for those contingencies and scenarios differ. Since forecasting tries to give insight into future developments it also has to present probabilities for the considered future developments in order to achieve a better basis for decision-making and planning. DeLurgio points out that probability in forecasts leads to results which comprise not only a single value, but a range of values. He mentions that “a good forecast statement is ‘expected sales next month for product X is 400 units, with a 70 percent probability that sales will be 300 to 500 units’” (DeLurgio 1998, p.10).
Furthermore, it is essential to base any forecasting activities on an appropriate reasoning system or methodology. Without such a reasoning system or methodology a forecast is less convincing and loses its objectivity and capability to decrease uncertainty, especially when it comes to business decisions. One can argue that a forecast, without any reasoning system as a logical base, degenerates into speculation, “a statement about a future condition in which the proponent conveys a high degree of uncertainty and/or lack of a rigorous, supportive rationale. The speculation openly rests on opinion, unproven hypothesis and imagination.” (Bright 1979, p. 234)

The discussion of the importance of horizon lengths or timing, probabilities, and reasoning systems or methodologies for forecasting allows the development of a general definition for forecasting as proposed by Bright:

“A statement specifying a future technological condition and its timing, arrived at through a system of reasoning consciously applied by the forecaster and exposed to the recipient.” (Bright 1979, p. 234)

This definition, however, does not cover all key aspects which were mentioned before. It has been shown that probabilities in forecasting are crucial for decision-makers in order to create wide-ranging contingency-plans. Moreover, since this thesis focuses on forecasting in strategic innovation management, probability becomes even more important. Decisions have to be made on the basis of a very long time frame and can lead to a great amount of expenses and investments. Therefore, decision-makers in the field of strategic innovation management face a greater amount of uncertainty and higher risks. For this reason, they have a greater need for probabilistic estimates of future scenarios and developments.

In order to achieve a comprehensive definition of forecasting, the definition of Bright has to be revised:

A probabilistic estimate of an uncertain future condition or development and its timing, arrived at through a system of reasoning consciously applied by the forecaster and exposed to the recipient.

2.2.2.2 Technology Forecasting

In the following section the general definition will be specified into a more precise definition of technology forecasting. Therefore, it is fundamental to understand the characteristics which are unique to this type of forecasting.
As Granger points out, technology forecasting evolved out of the argument that, in the long run, technological change is one of the most important influencing factors of economies. “One may expect that over periods of twenty years or more ahead the major changes in the economy and society will be due to changes in technology or in largely exogenous variables such as population size, distribution, or climate” (Granger 1989, p.209). Thus, technology forecasting seems to be most valuable when applied to long horizon lengths, which becomes even more important in case of strategic innovation management. For example, decisions for general strategic business planning are often based on a forecast horizon length of three to twenty years (DeLurgio 1998, p. 8).

Besides longer horizon lengths, the scope of the results is another specific property of technology forecasting. Such forecasts “are generally concerned with the characteristics of a technology rather than how these are achieved” (Granger 1989, p. 210). It was Bright who incorporated this fact into a definition of technology forecasting:

“Technology forecasting is a quantified statement of the timing, the character or the degree of change in technical parameters and attributes in the design, production and application of devices, materials and processes, arrived at through a specified system of reasoning.” (Bright 1979, p. 235)

The general definition of forecasting as presented in the previous paragraph and the fact that technology forecasting mostly deals with long horizon lengths, require a revision of Bright’s definition to attain a more rigorous and precise definition of technology forecasting:

Technology forecasting is a probabilistic, long-term estimate of the timing, the character or the degree of change in technical parameters and attributes in the design, production and application of devices, materials, and processes, arrived at through a system of reasoning consciously applied by the forecaster and exposed to the recipient.

2.2.2.3 Forecasting vs. Foresight

As introduced previously, forecasting is an activity to systematically achieve insight about future trends and developments. But in many occasions forecasting is confused with foresight and sometimes these terms are even used as synonyms. Still, there is a difference between these terms.
Martin defines foresight as

“The process involved in systematically attempting to look into the longer-term future of science, technology, the economy and society with the aim of identifying the areas of strategic research and the emerging generic technologies likely to yield the greatest economic and social benefits.”

(Martin 1995, p. 104)

The definition of Martin points out that foresight targets a very broad range of future developments to derive general conclusions and recommendations. In addition to that, lexicographically foresight means “The action or faculty of foreseeing what must happen”, “The action of looking forward”, and “Care or provision for the future”, among other meanings (Oxford English Dictionary, Second Edition 1989). In other words, foresight comprises aspects of forecasting as well as planning. The *Handbook of Knowledge Society Foresight* says that “Foresight can include both the activities of forecasting and planning. Planners can use forecasting methods to predict the outcomes of alternative plans. If the forecast outcomes are not satisfactory, planners can revise the plans, then obtain new forecasts, repeating the process until the outcomes are satisfactory. These kinds of processes are typical of pragmatically oriented foresight processes.” (Miles, Keenan, and Kaivo-Oja 2002, p. 163).

In general, forecasting “is related to the principle of ‘what, if’ or ‘what will happen, if’” (Miles, Keenan, and Kaivo-Oja 2002, p. 163), while foresight is a broader scenario based view of the future, with elements of planning and forecasting alike, which also tries to connect different areas of future developments, e.g., technology and society.

### 2.2.3 The Technology Forecasting Process

The technology forecasting process can vary in different situations from a relatively simple process with just a few stages to a process comprising a complex structure of stages and sub-processes (DeLurgio 1998, p. 26). For that reason, the basic framework common to all process forms will be introduced first, followed by a description of the technology forecasting process for the use within strategic innovation management, which will be the basis for the subsequent chapters of this thesis.
Figure 2-1 Stages of Forecasting

Figure 2-1 illustrates the six basic stages of forecasting as presented by Armstrong (Armstrong 2001, p. 8). These stages, in the same or a very similar order, can also be found in other literature (DeLurgio 1998, p. 27 or Reger 2001, p.538), sometimes in combination with additional stages. According to DeLurgio, the forecasting process is most effectively approached using the so-called *scientific method* which leads to the observed process structure (DeLurgio 1998, p. 26). In the following pages, each single stage will be characterized.

**Formulate Problem** – The first stage is a consequence of the situations to which forecasting is applied. It has been pointed out that forecasting is a tool for decision-support and planning. Hence, the problem situation has to be examined in detail and factors like objectives, target scope or time horizon have to be defined in order to make informed decisions. Reger, for example, mentions two major research areas which companies can target; “core technologies” and “white spaces” (Reger 2001, p. 539). Forecasting activities which target “core technologies” incorporate future developments and technologies which are specific to the companies’ domain, while a scope of “white spaces” uses a broader range of future developments to facilitate the identification of new research fields, technologies or radical innovations. When objectives, scope and time horizon have been defined, the information need can be derived which leads, once known, to the next stage.

**Obtain Information** – At first, the type of information which is needed and the corresponding sources are selected, which can either be formal (e.g. journals, reports, trend studies, etc.) or informal (e.g. workshops, conferences, personal contacts, etc.) (Reger 2001, p. 541). The main activity in this stage, then, is to collect information about the problem setting and its environment to be able to identify and better understand factors which influence the problem setting. Examples for these factors in a production planning scenario can be past data about sales of a product, out-of-stock conditions, prices, sales of competitors’ products, and the like (DeLurgio 1998, p. 26).
The information gathered helps to concretize and structure the formulated problem and prepares the following stage.

*Select Methods* – Now that the problem is formulated, influencing factors identified, and information about the environment of the problem collected, the methods that will be used to conduct the actual forecast are selected. The type of information collected in the previous stage is one criterion of method selection. Additional criteria include, but are not limited to, the objectives of the forecast, the targeted time horizon, the scope of the forecast, the availability of information or expertise, and the expected degree of technological change. Armstrong discusses method selection in greater detail in his book *Principles of Forecasting* and presents a selection tree for forecasting methods which guides forecasters during method selection (Armstrong 2001, p. 376). A general overview of existing forecasting methods will be given in section 2.2.4 of this thesis.

*Implement Methods* – The necessary activities with respect to the selected forecasting method or the set of combined methods are conducted within this stage, which forms the main part of the forecasting process and includes the highest amount of active participants in the process. Many forecasting methods incorporate external expertise in form of technology specialists, researchers, or other sources in addition to the companies’ R&D department or forecasting office. Within the scientific method this stage is called “Experiment execution” (DeLurgio 1998, p. 27).

*Evaluate Methods* – In order to achieve a successful and accurate forecast it is necessary to evaluate and assess the methods used and the corresponding results with regards to questions like:

- Do the results meet the objectives set during problem formulation?
- Has the targeted scope been achieved?
- Have the selected methods been appropriate for the targeted type of problem?

In the case that questions like those stated above cannot be answered to a satisfying extent, the method selection has to be revised and the implementation of the forecasting methods repeated. For information about the evaluation of forecasting methods the reader is referred to Armstrong’s *Principles of Forecasting*, pages 443 – 472.

*Use Forecasts* – This is the process of using the resulting forecasts to find solutions for the formulated problem. DeLurgio mentions that ongoing maintenance and verification
are necessary to ensure that the results are still valid and effective (DeLurgio 1998, p. 27). Hence, it is recommended to monitor reality and compare it to the forecasts in order to be able to respond to possible inaccuracies.

Since this thesis targets technology forecasting within the scope of strategic innovation management, the process will be rendered more precisely to comply with the business situation of R&D program planning and to support decision-making. In this context the mentioned suggestion of ongoing monitoring becomes even more important, since companies have to respond to changes as quickly as possible to stay competitive. In addition to that it can be assumed that in a large company individuals who are conducting the forecast and decision-makers are not the same people. Therefore, additional steps in order to prepare decisions and then make decisions are necessary for a complete view of the process. To include these thoughts into the process, the last stage of the process has to be split and a more detailed structure created. The resulting technology forecasting process for strategic innovation management is shown in figure 2-2.

![Figure 2-2 Technology Forecasting Process for Strategic Innovation Management](image)

### 2.2.4 Overview of Forecasting Methods

For the later discussion technology forecasting, it is important to get a basic understanding of available classes of forecasting methods. This subchapter is based on the “Methodology Tree” by Armstrong (Armstrong 2001, p. 9) which illustrates the characteristics of forecasting methods and their relationships. Figure 2-3 depicts the Methodology Tree.
Armstrong begins with a separation of judgmental and statistical methods. He mentions, however, that judgment pervades all aspects of forecasting (Armstrong 2001, p. 9). As further down a method is positioned in the tree as higher is the amount of the judgmental and statistical integration. On the judgmental side of the tree the methods are split into those predicting one’s own behavior and those predicting the behavior of others, mostly by including experts into the forecasting process. On the side of method types predicting one’s own behavior, the methods are characterized by the influence of a role. If a role influences the decision to make, role playing is a valuable tool for forecasting the outcome of the decision through the simulated interaction of roles affected by the decision. In case there is no influence of a role, the intentions method can be used in which people predict their own behavior in different situations. Conjoint analysis goes a step further than the intentions method by trying to create a connection between personal intentions and certain features of a situation through statistical analysis. For example, “a forecaster could show various designs for a computer and ask people about their intentions to purchase each version” (Armstrong 2001, p. 9).

Forecasting methods within the others branch base on expert opinions about how organizations or others will behave. There is a broad number of forecasting methods which belong to this type, with the Delphi method being the most famous one. In this method questionnaires are sent out to experts in the targeted fields who answer the
questions by the use of their subjective judgment. Once the questionnaires are sent back and analyzed, they are sent out again to the same experts together with the results of the first round in order to get a second estimation. The reason for this is to share the results and create a common knowledge base among all participants of the forecast. This process can be repeated for one or two more rounds after which the final conclusions are drawn and the forecast is created. Further information about the Delphi method can be found in the forecasting literature (for example: Armstrong 2001, DeLurgio 1998, Granger 1989, or Martino 1983). Judgmental bootstrapping refers to methods which use regression analysis in order to draw conclusions and rules out of expert opinions and, to a certain extent, belong to the class of expert systems.

Judgment and statistics are merged into one method type when analogies are used. Based on statistical data, experts try to forecast the development of a situation. The success of such an approach depends on the degree of similarity between the situation which has to be predicted and the one the statistical data is taken from.

The statistical side of the tree is split into univariate and multivariate methods. The univariate part of the tree contains extrapolation methods (Armstrong 2001, p.10), that is, values are predicted by the use of older values within a (time-) series. The simplest method of this type is using today’s number of sales to predict tomorrow’s number. When domain knowledge and knowledge about forecasting procedures is combined in a type of expert system to achieve this task, one speaks of rule-based forecasting. Full expert systems utilize an even greater integration of expert rules (rules which are similar to the way experts create their judgments) in order to support forecasting.

Multivariate forecasting methods are distinguished whether they are based on statistical data or theory. The latter leads to econometric models which base on domain knowledge or findings from prior research. “Econometric models provide an ideal way to integrate judgmental and statistical sources” (Armstrong 2001, p.10).

In general, Armstrong characterizes eleven different forecasting method types. For more detailed information on the different types and for method examples the reader is referred to the corresponding forecasting literature.
2.3 Knowledge Management

2.3.1 Motivation and History

Long before knowledge management became an issue, the nature of knowledge had already been discussed. The roots of knowledge go back to the days of the great philosophers. Literature always distinguishes between western and eastern intellectual tradition, but in both cases philosophers or clericalists were the first who started to explore the meaning of knowledge, although with a different focus and different influences through religion and society than observed today (Wiig 1999, p. 2). Nonaka and Takeuchi mention that the foundations of Western epistemology (the theory or science of the method or grounds of knowledge, Oxford English Dictionary, Second Edition 1989) go back to Plato and Aristotle. They write, “It was Plato who first built up an elaborate structure of thought on knowledge from a rationalistic perspective.” (Nonaka, Takeuchi 1995, p. 22)

Besides this philosophical approach toward knowledge, the practical need for knowledge or, more precisely, the need for practical expertise and understanding goes back even beyond the times of the philosophers mentioned above. Humans have always had a need for practical knowledge since the earliest fight for survival (Wiig 1999, p.2). A form of implicit unsystematic management of practical knowledge has always been tied to this need. As suggested by Wiig “the craft-guilds and apprentice-journeyman-master systems of the 13\textsuperscript{th} century were based on systematic and pragmatic knowledge management considerations.” (Wiig 1999, p. 2) Although the industrial revolution of the 17\textsuperscript{th} century had a great impact on the industrial landscape, the impact on the way knowledge was managed was rather low. This changed with the approach of the “knowledge era” in the 20\textsuperscript{th} century and its efforts to increase effectiveness. (Wiig 1999, p. 3)

In economic theories during the 20\textsuperscript{th} century varied interpretations for knowledge existed. Neoclassical economists partially ignored the influence of knowledge on a companies’ success and the existing markets and believed that “under market mechanism, every firm has the same fixed knowledge that enables profit maximization, rather than having different knowledge created by each firm”, while others “argued that knowledge is ‘subjective’ and cannot be treated as fixed” (Nonaka, Takeuchi 1995, p. 33).
History has shown that knowledge and its management as it is discussed today has many origins which can be summarized as follows:

- “Religion and Philosophy (e.g., epistemology) to understand the role and nature of knowledge and the permission of individuals ‘to think for themselves.’”
- “Psychology to understand the role of knowledge in human behavior.”
- “Economics and social sciences to understand the role of knowledge in society.”
- “Business Theory to understand work, and its organization.”
- “Rationalization of Work (Taylorism), Total Quality Management, and Management Sciences to improve effectiveness.”
- “Psychology, Cognitive Sciences, Artificial Intelligence (AI), and Learning Organization to learn faster than competition and provide foundation for making people more effective.” (Wiig 1999, p. 3)

### 2.3.2 Data, Information, and Knowledge

The preceding chapter has shown that knowledge and knowledge management have been approached from many different perspectives. Each perspective includes its own assumptions, criteria, and characterizations leading to differing definitions for terms used. In common language the terms *data, information*, and *knowledge*, especially the latter two, are very often used synonymously. Unfortunately, the knowledge management community has not been able to find universal definitions for these terms, although discussed for many years. Aamodt and Nygård argue that a “possible reason is that several perspectives easily get mixed in discussions about definitions of concepts that are polymorphic. ... In order to get the meaning of a polymorphic (non-classical) concept, it has to be understood within a particular context [19], i.e. related to some purpose or intended use, and seen from a certain perspective” (Aamodt, Nygård 1995, p. 3). Hence, no single, true definition exists for each of these terms. In fact, several definitions are equally valid when used within a specific context. In the next part of this chapter the terms *data, information*, and *knowledge* will be discussed and defined, while the reader is asked to keep the business orientation of this thesis as the focus of the developed definitions in mind.
2.3.2.1 Data

Data are commonly accepted as the smallest entities in the domain of knowledge management. Davenport and Prusak define data as

\[ \text{"a set of discrete, objective facts about events. In an organizational context, data is most usefully described as structured records of transactions."} \] (Davenport, Prusak 1998)

Additionally, they stress that data “says nothing about its own importance or irrelevance.” (Davenport, Prusak 1998). Taking this into account, data comprise characters, signals, patterns or signs which all can be characterized by having “no meaning for the system concerned” (Aamodt, Nygård 1995, p. 8). Thus, the most defining property of data is the lack of meaning. In other words, “data require minimal human judgment” (Tsoukas, Vladimirou 2001, p. 976). Moreover, since data do not include any meaning, they have to be interpreted in order to be valuable. This fact has been incorporated into the definition of data by Aamodt and Nygård which will be used within the course of this thesis:

\[ \text{"Data are syntactic entities - data are patterns with no meaning; they are input to an interpretation process, i.e. to the initial step of decision making."} \] (Aamodt, Nygård 1995, p. 6)

2.3.2.2 Information

Information is often described as a message or several messages. (Nonaka, Takeuchi 1995, p. 58; Davenport, Prusak 1998), which usually has (respectively have) a written, audible, or visible form. Like every message, information has a sender and a receiver, while the receiver is the one being informed. Due to that, it is the receiver who decides “whether the message he gets is really information -- that is, if it truly informs him” (Davenport, Prusak 1998). Nonaka and Takeuchi mention that “information provides a new point of view for interpreting events or objects, which makes visible previously invisible meanings or sheds light on unexpected connections” (Nonaka, Takeuchi 1995, p. 58). Obviously, it is important for the characterization of information that it includes meaning and is context-specific. This very meaning and hence, information itself is created through the transformation of a syntactic structure into a semantic, meaningful entity (Aamodt, Nygård 1995, p. 7). Aamodt and Nygård illustrate that a random sequence of symbols, such as U4$µ+?@24r+, “or a series of signals from a sensor, is data to most human beings, while the data items ‘inflation rate’, ‘decreased blood
pressure’, and ‘the Cuban crisis’ have meaning, and therefore are information” (Aamodt, Nygård 1995, p. 8). Consequently, they propose the following definition for the term information:

“Information is interpreted data - information is data with meaning”

(Aamodt, Nygård 1995, p. 6)

This definition complies with Davenport and Prusak, who express that “data becomes information when its creator adds meaning” (Davenport, Prusak 1998). It is the generic character of the stated definition which makes it applicable to, for example, small pieces of information as mentioned in the example earlier, as well as, whole books or internet web-pages. As a result, the stated definition of information will be used within the subsequent chapters of this thesis.

2.3.2.3 Knowledge

It has been shown that the term information can be used in a broad range of situations and that its main characteristic is meaning. Nonaka and Takeuchi argue that “knowledge, like information, is about meaning. It is context-specific and relational.” (Nonaka, Takeuchi 1995, p. 58). This raises the question of where exactly the difference between information and knowledge lies. McInerney answers that “unlike static information that can be held in databases and on paper, knowledge is based in sentient beings, or emanates from them, and thus, it is always changing with the human experience.” (McInerney 2002, pp. 1009-1010). That is why knowledge is not as tangible and concrete as information. Rather, it bases on a dynamic process within the human mind. Davenport and Prusak apply a similar approach and define knowledge as

“a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knowers.” (Davenport, Prusak 1998)

In addition to that, knowledge is much more far-reaching than information. That means “even when formed within a particular context and for a particular type of use, [knowledge] may be reused dynamically in future situations that are different. The 'tension' within a body of knowledge to be both specific and related to the use from which it originated, as well as being general and flexibly usable, is a feature that distinguishes it from data and information.” (Aamodt, Nygård 1995, p. 11). Besides the
focus on flexibility, Aamodt and Nygård include the application of knowledge in their perspective. As Nonaka and Takeuchi stress, knowledge is essentially related to human action (Nonaka, Takeuchi 1995, p. 59). It is, for example, the enabler for “the individual capability to draw distinctions, within a domain of action, based on an appreciation of context or theory, or both” (Tsoukas, Vladimirou 2001, p. 976).

In the context of this thesis, a definition of knowledge is chosen which is aligned to the argumentation of Alavi and Leidner:

“Knowledge is a justified personal belief that increases an individual’s capacity to take effective action. …[It] is information possessed in the mind of an individual: it is personalized or subjective information related to facts, procedures, concepts, interpretations, ideas, observations and judgments (which may or may not be unique, useful, accurate, or structurable).” (Alavi, Leidner 1999, pp. 6-7)

The specific feature of knowledge of being closely tied to the human mind is the reason for some researchers to propose a further refinement of knowledge. They distinguish between “tacit” which is also referred to “implicit” knowledge and “explicit” knowledge. Nonaka and Takeuchi, for example, adopt this distinction, which originally has been identified by Michael Polanyi, and state that “tacit knowledge is personal, context-specific, and therefore hard to formalize and communicate. Explicit or 'codified' knowledge, on the other hand, refers to knowledge that is transmittable in formal, systematic language.” (Nonaka, Takeuchi 1995, p. 59) In other words, tacit knowledge is strictly linked to the human mind, while explicit knowledge can be extracted into knowledge representations (e.g. documents, books, etc.) which can be transmitted in order to become input, and thus information, for another knowledge creation process.

In addition to the presented definition of knowledge, this distinction between tacit and explicit knowledge will also be used within the thesis in order to provide a more precise perspective where necessary.

2.3.3 Definition of Knowledge Management

Knowledge management has not always been a topic within companies, as history has revealed. Nevertheless, with the arrival of the “knowledge era” and the identification of knowledge as a foundation for competitive advantage, companies have started to introduce knowledge oriented processes and activities in order to establish knowledge
management. Von Krogh, Ichijo, and Nonaka criticize that knowledge management is not integrated into the companies’ competitive strategies, but rather, “‘the knowledge issue’ tends to become the responsibility of human resources, information technology groups, or corporate R&D; sometimes it is only part of initiatives located deep within various business units” (von Krogh, Ichijo, Nonaka 2000, p. 71). A possible reason might be the confusion with the term knowledge management, which is often used although information management is meant. Comparable to the relation of information and knowledge, knowledge management is also much more far-reaching than information management. Newman and Conrad define:

“Knowledge management is a discipline that seeks to improve the performance of individuals and organizations by maintaining and leveraging the present and future value of knowledge assets.” (Newman, Conrad 1999, p.1)

The authors focus the improvement of the performance of individuals and organizations with knowledge being the medium, whereas knowledge itself is the central aspect of McInerney’s definition of knowledge management:

“Knowledge management (KM) is an effort to increase useful knowledge within the organization. Ways to do this include encouraging communication, offering opportunities to learn, and promoting the sharing of appropriate knowledge artifacts.” (McInerney 2002, p. 1014)

Both definitions stress important characteristics of knowledge management. At first, knowledge is obviously the focus of any process which belongs to knowledge management. Furthermore, the goal of effective knowledge management is to improve the performance of the company and its individuals through support of knowledge creation, dissemination, and utilization. Alavi and Leidner include these views into their definition of knowledge management and also take the importance of the support for both tacit and explicit knowledge into account which is the reason for the adoption of the definition for this thesis:

“Knowledge management ... refers to a systemic and organizationally specified process for acquiring, organizing and communicating both tacit and explicit knowledge of employees so that other employees may make use of it to be more effective and productive in their work.” (Alavi, Leidner 1999, p.2)
2.3.4 **Organizational Knowledge Creation**

A perspective to knowledge management which has had great influence on other researchers is the process of organizational knowledge creation. Nonaka and Takeuchi argue that an “organization cannot create knowledge without individuals.” Rather, it “supports creative individuals or provides context for them to create knowledge” (Nonaka, Takeuchi 1995, p. 59). Therefore, organizational knowledge creation can be seen as a knowledge management model which has a focus that complies with the given definition of knowledge management: “knowledge and individuals”, while it is the company’s task to ensure a supportive environment. The authors continue that “organizational knowledge creation … [is] a process that ‘organizationally’ amplifies the knowledge created by individuals and crystallizes it as a part of the knowledge network of the organization” (Nonaka, Takeuchi 1995, p. 59). As a consequence they propose the “Four Modes of Knowledge Conversion” or SECI-Model (Socialization, Externalization, Combination, and Internalization) as a model for organizational knowledge creation.

![Figure 2-4 The SECI Process](Adapted from Nonaka, Takeuchi 1995)

Figure 2-4 illustrates the four modes of knowledge conversion which will be described briefly within the rest of this section to facilitate understanding of the process (based on Nonaka, Takeuchi 1995 and Nonaka, Toyama, and Konno 2001).

**Socialization** – This is the process of creating and sharing tacit knowledge through shared experiences. Social interactions in, for example, informal meetings support the
creation of mutual trust and allow an exchange of tacit knowledge like perspectives and mental models. This process occurs within as well as beyond company boundaries.

*Externalization* – This is the process of extracting tacit knowledge into explicit knowledge which then can be codified and, thus, shared with others. The development of product concepts or, in general, the formulation and documentation of an idea are examples for this process.

*Combination* – Within this process explicit knowledge is used to create new explicit knowledge through collecting, editing, combining, or processing the explicit knowledge in another way.

*Internalization* – The process of transforming explicit knowledge into tacit knowledge and incorporating it into an individual’s tacit knowledge base is called internalization, with all activities taking place in the individual’s mind. This is achieved through reflection, interpretation or all different forms of learning of explicit knowledge. The newly adopted tacit knowledge can then be used to initiate another cycle of knowledge creation through socialization.

Obviously, this model is just one example of approaching organizational knowledge creation and knowledge management. Comparable to the definitions of knowledge, other models exist which focus other aspects of knowledge management. While Nonaka and Takeuchi concentrate on the creation of knowledge through conversions between tacit and explicit knowledge, Probst, for example, proposes the concept named “Eight Components of Knowledge Management”, which approaches knowledge management from a management perspective following the “St. Galler Management Concept” (Probst 2001, p. 254).

### 3 Analysis of Technology Forecasting

The previous chapter introduced the fields of strategic innovation management, technology forecasting, and knowledge management and defined the most important terms and concepts. The presented information provides the basis for this chapter which explores the problem setting to a greater extent and leads to the analysis of technology forecasting with respect to knowledge management. First, the need for knowledge management within technology forecasting for strategic innovation management is identified, after which the technology forecasting process is analyzed with respect to the
usage of information and knowledge and the potential of improved performance through supporting knowledge management technologies.

### 3.1 Knowledge Management Needs within Technology Forecasting

There are two perspectives which have to be considered in order to determine the need for knowledge management within technology forecasting. First, the topic can be approached from an inside-the-company view: which technological, organizational, or social factors within a company make it necessary to integrate knowledge management into the technology forecasting process. On the other hand, a company’s forecasting process is obviously influenced by the company’s environment, because most of the information which is collected during a forecasting process comes from outside the organization. Therefore, an analysis from an outside-the-company perspective is also crucial to achieve a complete view of the need for knowledge management within technology forecasting.

#### 3.1.1 Needs Emerging from Inside the Company

Inside a company, technology forecasting is closely linked with decision-making processes. In the context of this thesis, for example, it is part of the activities incorporated in strategic innovation management in order to support planning of innovation and R&D programs. DeLurgio argues that “it is important to recognize the role of forecasting in expanding the knowledge base of organizations and whole societies.” (DeLurgio 1998, p. 6). Thus, technology forecasting itself can be seen as a knowledge-creating activity; that is, knowledge in the sense of enabling managers to make strategic decisions and plan a technological innovation path for the company. Therefore, decision-makers need an as comprehensive view of future developments as possible, which cannot be achieved with the help of technology forecasting alone. The end product of forecasting activities is, in most cases, a report which represents all future developments analyzed. However, it can be assumed that this report does not contain enough information for a decision maker to recreate all the knowledge that has been created by participants through the whole forecasting process. Tacit knowledge, like perspectives and prior experiences shared by forecasters, might be valuable for a decision-maker. This facilitates interpretation of the information contained within the reports in a more efficient and comprehensive fashion, thus, leading to decreased uncertainty and better informed decisions. Moreover, reports cannot contain all information which is available for the forecasters. In order to provide precise
information and to reduce complexity of the document, some information has to be left out. However, this information might become useful at a later point of the decision process. Without efficient ways to recover the missing information, the decision process is either slowed down, due to the additional time spent to analyze or acquire the missing information for a second time, or it becomes less accurate.

Another aspect of technology forecasting creates the additional need for an integration of knowledge management into technology forecasting, particularly for knowledge management technologies. It has been shown that technology forecasting is a process that is based on information and knowledge and can also be seen as a knowledge-creating process itself. Furthermore, many participants possessing different roles are involved. Figure 3-1 illustrates the roles involved in the different steps of the technology forecasting process for strategic innovation management as presented in chapter 2.2.3.

![Figure 3-1 Roles within Technology Forecasting](Adapted from Miles, Keenan, Kaivo-Oja 2002, p. 163)

The varying involvement of the different roles shown above leads to a dynamic project team size which can change at almost every step of the process. Moreover, a social and organizational network as dynamic as the team size is created and has to be managed. This includes the communication and distribution of specific information and knowledge needed by individuals incorporating one of the roles in order to fulfill their tasks.

In summary, from an inside-the-company perspective four major reasons can be identified for the emerging need of knowledge management support for technology forecasting within strategic innovation management. Moreover, the last three reasons have the potential of improving the quality and efficiency of the process:
Technology Forecasting is a knowledge-creating process itself.

- Tacit knowledge which has been created during the process is not transferred to decision-makers.
- Resulting documents of the forecast can only deliver a static and limited amount of information and explicit knowledge.
- Technology Forecasting takes place within a team of dynamic size and structure in which information and knowledge has to be communicated and distributed.

### 3.1.2 Needs Emerging from Outside the Company

As already mentioned in the introduction of the thesis, making the right decisions with respect to future developments and technologies is vital for a company’s competitiveness. One reason for this is the decreasing length of technological lifecycles as “technological change is one of the most important forces affecting a firm’s competitive position” (Burgelman, Maidique, Wheelwright 1996, p. 6). Additional dynamics and uncertainty are created by the phenomenon of unexpected, disruptive innovations, which a company has to cope with and which can never be fully excluded. Another factor, which increases the pressure felt by decision-makers, is cost. Vanston states that “under pressure to contain these [higher] costs, it has become increasingly important for R&D programs to focus on projects that will result in enhanced profits and sustainable competitive advantage.” (Vanston 1996, p. 57). All these factors are evidence for how crucial it is for a company to make the right decisions in a constantly decreasing time frame.

On the other hand, the same reasons lead companies to face increasing uncertainty with respect to future developments. In order to deal with such uncertainty, companies have to collect and assess more information in a faster and more efficient way than they used to. This is also true in the context of technology forecasting within strategic innovation management. It can be assumed that more information leads to a reduced uncertainty and thus to better informed decision. At the same time, however, more information also leads to higher complexity and consequently to a decrease in efficiency and a slower process. The amount of information required and the decrease of uncertainty and the time needed to collect, assess and process information are anti-proportional to each other. Thus, companies have to find an equilibrium as much on the information side as
possible in order to keep uncertainty low, while keeping complexity on a level which can still be handled by the forecaster.

Hence, two main factors emerging from outside the company can be named which are related to a technology forecasting process’s efficiency, influence a company’s competitive advantage, and create a need for knowledge management within technology forecasting:

- Decisions have to be made faster to stay ahead of competition
- More information with an increasing complex relational structure has to be collected, assessed, and processed to decrease uncertainty

3.2 Analysis of the Technology Forecasting Process

Since the focus of the thesis lies on the evaluation of knowledge management technologies for the support of technology forecasting in strategic innovation management, it is the goal of this subchapter to analyze each step of the technology forecasting process, as described in chapter 2.2.3, and identify those steps capable of being supported by knowledge management technologies. Therefore, the information and knowledge used within a process step will be examined and the potential of improvement of the step determined.

3.2.1 Step I: Formulate Problem

As described previously, this step incorporates the definition of the objectives, the time horizon, and the scope of the technology forecasting activities. In the context of strategic innovation management the objectives of forecasting activities are most likely to forecast future technology innovation paths either to provide a basis for the creation of completely new sets of innovation strategies or to evaluate the current strategies by comparing them to the predicted future. In order to determine the time horizon and the scope of the forecast, the individuals responsible for this step have to utilize various pieces of information and knowledge.

The time horizon, as well as the scope of the forecast, depends on strategic and tactical goals for innovation and R&D which are tied to the strategic goals of the company. These goals are codified in documents or reports of some sort which are one source of information for this first step of technology forecasting. Additional sources of information can be taken from the current R&D project portfolio, sales reports, market
reports, and the like. The information which can be acquired through such sources is helpful in order to determine the gap between where the company stands and where strategy wants it to be, which might be especially valuable while determining the scope of the forecast. In addition to these hard facts, the activities of this step are influenced mainly by two different knowledge areas. One of them is directly connected to the individuals responsible for this step as their knowledge about forecasting in general and within their specific company has obviously great influence on the determination of a reasonable time horizon and scope and hence, on the efficiency of this first step. Such knowledge has been created, for example, through experiences during prior forecasting activities. A forecasting team that executes these tasks on a regular basis does not need much time for the definition of a reasonable time horizon and is able to create the scope of the forecast more precisely, because it is able to reuse or revise definitions from previous forecasts, while an inexperienced team has to create such knowledge before it can define the necessary components of the forecast. Secondly, all members of the same company share the same corporate culture which is aligned to the corporate strategy and the overall goals of the company. This leads to the conclusion that besides the codified strategic goals, each individual has a personal understanding of the future path of the company.

It has been shown that different pieces of information and knowledge have to be taken into account when examining factors influencing the first step and its efficiency. The character of the activities of this step raises the need for a collaborative environment which allows sharing of prior experiences and company strategies in order to work on the problem definition. In this context personal communication is a crucial factor and thus, the potential for an efficiency increase through the utilization of knowledge management technologies is rather low.

3.2.2 Step II: Obtain Information

As described in chapter 2.2.3, information of various kinds is collected during this process step in order to structure the formulated problem and render it more precisely. In the case of strategic innovation management, the main task of this process step is to find out which information is available, how it can be acquired, and whether it is valuable for the achievement of the objectives. This facilitates the selection of forecasting methods within the next step and also enables the development of an initial rough concept of topics and technologies which should be addressed by the forecasting
methods. Thus, the scope of the forecasting activities can be evaluated and revised if necessary.

One factor which influences the efficiency of this step is, as in the first step, the experience of the individuals who are involved in the process. Experienced forecasters have the knowledge about which sources and types of information have been appropriate in prior technology forecasts to achieve the defined objectives and thus, are able to make decisions about the quality of information and its sources faster than inexperienced individuals. On the other hand, technology forecasting, as explained previously, deals with the uncertainty and dynamics of technological progress. Hence, there is always a need for new information about the current business situation and the areas within the forecasting scope. This leads to the thought that the knowledge and experience of the forecasters do not make information collection an unnecessary activity of this process step. Therefore, the complexity of finding valuable information in the large amount of information available has to be reduced which raises the need for a supporting technology.

Knowledge is an enabling factor for the evaluation and assessment of the collected information. However, the dynamic and uncertain character of technological progress makes it necessary that such knowledge is kept on an up-to-date level. Therefore, a technology which supports the collection of information should also enable individuals to put that information into context and facilitate the creation and dissemination of new knowledge with respect to the scope of the forecast which leads to a high potential of improvement through knowledge management technologies.

3.2.3 Step III: Select Methods

The activities of this process step and their corresponding inputs depend on the results of the previous two steps. The collected information and the created knowledge about the current situation and the scope of the forecast from the second step are combined with the objectives defined in the first step of the process in order to develop criteria for the selection of appropriate forecasting methods. Clearly, methods that have been used before are preferred selections and the need for the creation of new knowledge is relatively low. Only when a set of technology forecasting methods did not deliver results which accomplished the objectives of the forecast and new and unknown methods have to be selected, must new knowledge be created. Such knowledge,
however, has to be formed through the internalization of external expertise or through other means of learning the new methods.

### 3.2.4 Step IV: Implement Methods

Obviously, the types of activities executed within process step IV depend on the selected set of forecasting methods and thus also the information and knowledge used within those activities depend on the selection.

One purpose of technology forecasting is the examination of the timing of specific technological innovations and breakthroughs. A method developed in 1963 to accomplish this task is the Delphi method which has already been described in chapter 2.2.4. Since the company is more or less only coordinating the survey and the distribution of the results to the participating experts, this method can be seen as an individual sub-process of the overall forecasting process. This sub-process has a defined forecasting scope as its input and statistical statements about the timing of innovations and breakthroughs as its output which is executed to a great extent outside the company. Thus, information technology in general can be one means of facilitating the execution of a Delphi analysis, but specific knowledge management technologies inside the company have no influence on the process. Thus, the potential support of a technology forecasting method through knowledge management technologies supposedly depends on the degree of a method’s execution inside the company which is conducting the forecast, which is rather small for the Delphi method.

A forecasting method which is entirely executed within the company is Scenario-Forecasting. This is especially suitable in case of strategic innovation management. “Starting with some specific situation, such as the observed present, a scenario attempts to set up a logical sequence of events to indicate how some future state or situation can evolve step by step.” (Granger 1989, p. 224). Each of those events within a scenario is characterized by a probability-estimation. Additional scenarios are created through the consideration of estimated disruptions, breakthroughs or similar events with an uncertain outcome. In such cases, several branches are created to each of which a certain probability is assigned. The advantage of this method is the application of relations between different technological areas and the resulting influences. In addition to those benefits Granger mentions that “the primary purpose of these scenarios is not strictly to predict the future but rather to facilitate a systematic exploration of branches or networks of critical events within some explicit time frame” (Granger 1989, p. 224).
It is obvious that such an approach to forecasting is highly complex and a large amount of information from different technological areas has to be assessed, reflected, and combined in order to create the knowledge necessary to develop reasonable scenarios. All of these tasks can be executed within a company by a single individual or more likely by a team of experts with different areas of expertise. The author assumes that such a situation has the highest potential of achieving an improved efficiency by the integration of knowledge management technologies.

In general it can be summarized that the extent with which knowledge management technologies are capable of supporting the selected forecasting methods depends on the degree of the internal execution of a method. The more activities of a forecasting method are conducted within the company the higher the need for information collection and knowledge creation and dissemination and hence, the higher the potential for a supporting technology to improve efficiency.

3.2.5 Step V: Evaluate Methods

As described in chapter 2.2.3 the purpose of this step is to evaluate the results of the implementation of the selected set of forecasting methods. The results are checked and reviewed with respect to the forecast’s objectives, scope, and time horizon. In case the objectives are not met, the third step is executed again and a different technology forecasting method is selected and implemented. Hence, there is no new information or knowledge used or created within this step. Even though the activities of this step are executed by a team, the structure of this step is rather simple which leads to the conclusion that there is no need for a specific support by knowledge management technologies.

3.2.6 Step VI: Prepare Decisions

This step incorporates the activities necessary to transform the results of the forecast into a set of innovation and R&D strategies and plans. The knowledge acquired during the previous steps has to be communicated to the R&D management which has not been involved in step II to step V. In addition to that it is useful to be able to recover and review specific pieces of information dynamically in case additional questions appear during application of the forecasting results and strategy creation. The availability of such information, however, depends on the selected forecasting methods to a certain extent. As shown previously, the roots of a result are easier to recover when a method
has been chosen with a high percentage of its activities being conducted inside the company and by company members. In such a case, the ability to explore the results and create a more comprehensive understanding of their origin can be a valuable way of improving the transfer of knowledge from the forecasters to the R&D managers.

### 3.2.7 Step VII: Make Decisions

The purpose of process step VII is to inform the responsible decision-makers about the developed strategies and plans. These persons can then decide about the implementation of these strategies and plans. This can either be top level management, members of the board, or the R&D management itself, depending on the objectives and the initiators of the forecasting activities. While the R&D management had to develop a certain amount of knowledge in order to use the outcome of the technology forecasting activities for strategy development and planning purposes, top level managers are concerned about how the proposed innovation and R&D strategy fits into the company’s competitive strategy. Thus, decisions of top level managers require a less comprehensive knowledge about the results and their origins than those of R&D managers. However, these decisions do require a combination of such knowledge with the knowledge about all the other strategic components of the company. Thus an integration of knowledge management technologies which support the technology forecasting process into an overall knowledge management concept might lead to an additional value and an increased efficiency for top level manager. This field, however, goes beyond the scope of this thesis and can be considered a field for further research.

### 3.2.8 Step VIII: Ongoing Monitoring

After implementing the strategies and plans developed during the decision preparation, it is the objective of step VIII to monitor whether the forecasts turn out to be correct or not. This ongoing monitoring is especially vital within strategic innovation management as explained previously. In order to implement such a monitoring it is necessary to sustain an active information flow into the company and to combine this new information with the knowledge which has been acquired through the earlier part of the forecasting process and, hence, keep it up-to-date. It is obvious that such a task might become quite complex over time as developments within different technology areas might indicate changes with an influence on the chosen innovation path of the company. It is the objective of this step to identify and observe these indicators and their relations with the technologies within the scope of the company, to initialize new forecasting
activities when the further progress of these technologies drifts away from the results of the previous forecast, and to allow the development of preemptive plans for possible contingencies.

In order to achieve the objectives of this process step a broad range of information has to be analyzed to identify possible relations of different research fields with technologies within the scope of the company. Furthermore, this information has to be scanned for potential precursors which indicate a disruptive innovation or breakthrough with an influence on the company’s innovation path. The results of these activities are input for the process of keeping the knowledge which has been acquired through the prior steps of the technology forecasting process up-to-date which then has to be disseminated and shared among members of the R&D management. Clearly, the execution of this task encompasses a very high complexity and can be assumed to be slow, cumbersome, and nearly insurmountable without technological support. This fact leads to a presumed high potential of efficiency improvement through the support of knowledge management technologies.

3.3 Summary

The main statement which is drawn from the previous sections is that technology forecasting efficiency can be increased through integrating and improving knowledge management. It has been emphasized that reasons for a general need for knowledge management are found inside a company, like the dynamic size of a forecasting team, as well as outside a company, like the increasing amount and complexity of available information. However, this need is not distributed evenly over the technology forecasting process. Rather, certain process steps, i.e. step II, step IV, step VI, and step VIII, appear to have a higher potential to be improved by the integration of knowledge management than others. Therefore, especially the support of these process steps leads to an increased efficiency of the whole technology forecasting process.

4 Evaluation of Knowledge Management Technologies

At this point of the thesis the reader should have a clear understanding about why a need exists for knowledge management support within technology forecasting for strategic innovation management. Moreover, awareness has been created for the potential increase of efficiency within each step of the technology forecasting process.
This chapter goes a step further and asks whether specific knowledge management technologies are capable of improving the problematic situation stated in the third chapter and why the application of such technologies can lead to the desired increase in efficiency. Therefore a characterization scheme is developed in which a knowledge management dimension is combined with a technology forecasting dimension. The resulting matrix enables the distinction of knowledge management technologies with respect to their value for the forecasting process and will be applied in the following analysis of selected knowledge management technologies.

4.1 Development of a Characterization Scheme for the Evaluation

To be able to evaluate and delineate knowledge management technologies a characterization scheme will be developed in the subsequent sections. Since such technologies differ with respect to knowledge management as well as technology forecasting, the scheme will combine the two fields through the integration of a dimension for both of them.

4.1.1 The Technology Forecasting Dimension

The analysis of the technology forecasting process as presented in the third chapter provides a good starting point for the development of a dimension for the technology forecasting perspective. One argument of the analysis of the need for knowledge management support is the fact that technology forecasting itself is a knowledge creating process. A second look at the forecasting process reveals that each step can be seen as a transformation process with specific inputs and outputs. The second process step, “Obtain Information”, for example, needs the definition of the forecasting objectives, the scope, and the time horizon as inputs. This information is utilized within the activities of the process step and thus transformed into information of a higher complexity through the combination of the input information with new information. New relations between certain pieces of information are identified leading to a greater complexity of the observed information structure. The subsequent step III, “Select Methods”, also requires input from the preceding steps. It is, however, different from step II with respect to the transformation of information. While the activities of process step II increase the overall complexity of the information structure, the complexity remains at the same level during step III, because the information is only analyzed to select suitable forecasting methods. An analysis of the other process steps shows that the steps of the technology forecasting process can be characterized by their degree of
increasing complexity; in other words, either the level of complexity is raised or it remains unaltered. Figure 4-1 illustrates this relation on an abstract level without claiming to represent the actual degree of complexity increase.

![Figure 4-1 Information-Structure Complexity over the Technology Forecasting Process](image)

When examining figure 4-1 it is possible to identify four steps which cause an increasing complexity of the information structure within the forecasting process. These steps are “Obtain Information”, “Implement Methods”, “Prepare Decisions”, and “Ongoing Monitoring”. Reasons for this attribute of these four steps can be found when analyzing the activities of each step. They all have the combination of results from previous steps with newly acquired information in common which leads to the creation of new knowledge. Such knowledge is needed to complete the tasks of each process step. Within step VI, “Prepare Decisions”, for example, the results of the implemented forecasting methods from step IV are used as input information and combined with the knowledge about the competitive strategy and capabilities of the company to create the necessary knowledge in order to develop appropriate alternatives for the company’s innovation strategy.

Another striking point of the analysis of the technology forecasting process is the fact that all these steps have been identified to incorporate a high potential for the support through knowledge management technologies. The attribute of increasing complexity, as stated above, might be an additional component of the explanation for these findings.

As a result of the development of a dimension for the technology forecasting perspective, the level of the information structure complexity is chosen which is expressed by the four process steps identified. A dimension of such a kind enables the
categorization of the knowledge management technologies by their capability to support these four process steps and allows an implicit visualization of the level of complexity a knowledge management technology can handle.

### 4.1.2 The Knowledge Management Dimension

While the development of the technology forecasting dimension can be based on the analysis of the forecasting process, a different approach has to be found to define the knowledge management dimension. As a starting point the definitions of data, information, and knowledge should be reconsidered. Since a defined difference between these terms exists, one can argue that the definitions of data, information, and knowledge could be used as a structure to categorize knowledge management technologies, e.g. the category *information* contains all those technologies which target information. Furthermore, the development of these definitions has shown that transformation processes are required to turn data into information and information into knowledge. A categorization structure which is based only on the definitions of the three terms is not capable of integrating such transformation processes and it is obvious that knowledge management technologies exist, which, for example, specifically support the transformation of data into information. Aamodt and Nygård propose a model for data, information, and knowledge which takes the relationships of these three terms into account (Aamodt, Nygård 1995, p. 8). Figure 4-2 illustrates this model.

![The Data-Information-Knowledge Model](image)

*Figure 4-2 The Data-Information-Knowledge Model (Aamodt, Nygård 1995, p. 8)*

The model explains the mentioned processes which are needed to transform, for example, data into information in addition to the basic structure of data, information, and knowledge. Moreover, it includes the authors’ perspective on the influence of knowledge on the transformation processes. With respect to the development of a
dimension for the characterization of knowledge management technologies from a knowledge management point of view, however, it can be argued that this model is not applicable. As it has been explained during the definition of knowledge, knowledge is closely linked with human action and the human mind with learning being one way of creating knowledge. While there might be a number of knowledge management technologies which support learning, there are no technologies which target knowledge itself as it is defined in this thesis.

Another disadvantage of such a model is the granularity. It can be assumed that several types of knowledge management technologies exist, which target information but with a different focus or different application areas. Therefore, a finer granularity is needed which, in an optimal case, can be based on a single and continuous criterion to facilitate adoption and the development of a knowledge management dimension for the characterization scheme as stated above.

Smolnik, Kremer, and Kolbe suggest an approach which fulfils the mentioned requirements and is based on the importance of context, called “The Continuum of Context Explication” with context explication meaning “discovering implicit meanings and expressing those meanings explicitly” (Smolnik, Kremer, Kolbe 2005, p. 28). The authors stress that context is an important aspect which many definitions of knowledge have in common (Smolnik, Kremer, Kolbe 2005, p. 30). Furthermore they compare several definitions of context. Dey and Abowd, for example, define context as follows:

“Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.” (Dey, Abowd 2000, pp. 3-4)

Besides the definition of knowledge, context also plays an important role in the definition of information. Nonaka and Takeuchi argue that “knowledge, like information, is … context-specific and relational.” (Nonaka, Takeuchi 1995, p. 58) Smolnik, Kremer, and Kolbe found that knowledge management technologies “focus on contextual information in different ways and with varying intensity” (Smolnik, Kremer, Kolbe 2005, p. 36). Consequently, the authors present five approaches to “find and use information objects and contextual information …, each with a differing degree of context and explication ease” (Smolnik, Kremer, Kolbe 2005, p. 36).
In the following paragraphs, each approach identified by Smolnik, Kremer, and Kolbe is described briefly (Smolnik, Kremer, Kolbe 2005, pp. 37-39):

*Data Approach* – As shown during the definition of the term *data*, data are symbols or signs without a meaning. The authors add that a context does not exist in the case of data and thus cannot be explicited. Nevertheless, several methods can be applied to transform data into information or even into domain-specific knowledge. The data approach encompasses these methods.

*Information Approach* – Most important for the definition of information is that it includes meaning. Additionally, it is obvious that each piece of information has a specific context. However, Smolnik, Kremer, and Kolbe argue that the “context is … interwoven with the content and difficult to conceptualize, which means that the methods implemented to find requested information objects have to rely on the content and cannot access contextual information. … No effort is therefore made to explicate context, as neither the authors nor the users provide or use explicit contextual information.” (Smolnik, Kremer, Kolbe 2005, p. 37).
Descriptor Approach – The addition of explicit contextual information to pieces of information and thereby providing context aware methods for information search and discovery is called a descriptor approach. Hence, this approach comprises methods which apply metadata in order to identify context specific information. Such methods require some effort to enrich information with explicit contextual information as authors have to provide necessary metadata at the time of creation.

Meta-context Approach – This approach extends the descriptor approach as explicit contextual information no longer resides only within the information but is integrated into a meta-layer of contextual information which lies above and spans a variety of information. Within this meta-layer different explicit contextual information and thus different context is connected. This leads to users being “enabled to discover new contexts and to leverage and enhance their knowledge.” (Smolnik, Kremer, Kolbe 2005, p. 39). On the other hand, a higher effort is necessary to create and maintain such a meta-layer.

Knowledge Approach – The knowledge approach focuses the human being and considers the tacit aspects of knowledge, e.g. competencies, experiences, values, and insights, which create a person-specific context in addition to the contextual information used in the previous approaches. This approach is about knowledge creation through actions like communication, construction, or cognition.

The continuum’s consideration of context and its explication offers a continuous criterion with which it is possible to distinguish different knowledge management technologies. This fact makes the Continuum of Context Explication an ideal basis for the development of a knowledge management dimension. Each approach forms one category which can be used to classify knowledge management technologies. The only exception is the knowledge approach. Since the knowledge approach is tied closely with the human mind and human action, knowledge management technologies cannot explicate person-specific context and thus this approach is not used within the knowledge management dimension.

4.1.3 The Context-Complexity-Matrix

The combination of the developed technology forecasting dimension with the dimension for the knowledge management perspective results in the creation of the Context-Complexity-Matrix as illustrated by figure 4-4.
Figure 4-4 The Context-Complexity Matrix

Such a matrix allows the characterization of knowledge management technologies with respect to the degree of context as well as the degree of the information structure complexity of the technology forecasting process. The background of each dimension implicitly provides further characteristics of the classified knowledge management technologies. A classification of Meta-context Approach/Step VIII – Ongoing Monitoring, for example, explicitly means that the knowledge management technology is capable of supporting the high complexity of step eight and, thus, the step itself and also comprises a high level of context. As described previously, this classification also implies that a certain effort is needed to implement and maintain such a technology as described previously. The matrix is applied to characterize knowledge management technologies during their evaluation in the rest of this chapter.

4.2 Evaluation

Now that a characterization scheme has been developed, it will be applied to the evaluation of selected knowledge management technologies within the following sections.

As mentioned previously, what is called knowledge management technology within this thesis is sometimes also referred to as knowledge management method or knowledge
management instrument. In the context of this thesis the term knowledge management technology is defined as

*a concept which can be implemented into a computer system with functionalities capable of supporting or automating knowledge management tasks.*

The breadth of available knowledge management technologies ranges from very simple to very complex ones. Therefore, the set of knowledge management technologies for the following evaluation has been selected to represent this breadth, namely: data mining, case-based reasoning, information retrieval, topic maps, and ontologies. Additionally, these knowledge management technologies seem to be most widely known.

**4.2.1 Data Mining**

**4.2.1.1 Introduction and Definition of Important Terms**

Data mining is often referred to as the natural evolution of information and database technology. An evolutionary path including the functionalities data collection, data access, data navigation, and data mining has been described by different authors in this or a similar form (e.g. Krahl, Windheuser, Zick 1998, pp. 25-26; Han, Kamber 2001, p. 1).

*Data collection* became possible with the appearance of the first computers and dealt with questions like “What has been the average revenue of the last five years?” The data to answer such questions were stored on streamer drives or similar devices which made a certain effort necessary to collect the appropriate data. The introduction of databases improved the availability of data and enabled *data access*. More sophisticated queries were possible through languages like Structured Query Language (SQL) which also led to an increased performance of data transactions. Following that, *data navigation* facilitated the analysis of data through concepts like Online Analytical Processing (OLAP). OLAP enabled users to change the focus of a data analysis interactively through data summarization, consolidation, and aggregation. It allowed users to view, for example, sales data summarized by city and then apply commands like “roll-up” to decrease granularity, meaning the data to analyze was then presented summarized by country.

The progress of the evolution and, hence, the development of *data mining* at the end of the last century benefited from several trends, e.g. the exponential growth of data and
information repositories through electronic checkout systems and the like, the implementation of data warehouse systems, and the ongoing performance increase of today’s computers and storage devices (Krahl, Windheuser, Zick 1998, p. 25). This development still leads to an increasing complexity of data analysis making it insurmountable for an individual without more sophisticated methods and tools. Han and Kambler write, “As a result, data collected in large databases become ‘data tombs’ - data archives that are seldom visited. Consequently, important decisions are often made based not on the information-rich data stored in databases but rather on a decision maker's intuition, simply because the decision maker does not have the tools to extract the valuable knowledge embedded in the vast amounts of data.”¹ (Han, Kambler 2001, p. 4). Accordingly, this large amount of available data and the need for transforming it into information has been identified to be one of the major reasons for the increasing growth of data mining and its application (Han, Kambler 2001, p.1).

As explained above, simple database queries and concepts like OLAP were predecessors of the domain of data mining. Lusti argues that the aim of data mining is the automated and non-trivial search for knowledge in mass-data. The search methods are non-trivial, because they utilize complex methods from the domains of knowledge-based systems and statistics, instead of common database tools (Lusti 2002, p. 260). According to Han and Kambler, “data mining involves an integration of techniques from multiple disciplines such as database technology, statistics, machine learning, high-performance computing, pattern recognition, neural networks, data visualization, information retrieval, image and signal processing, and spatial data analysis.” (Han, Kambler 2001, p. 9). In both cases the authors stress that data mining combines techniques from a variety of different domains to offer more sophisticated methods than existed before. The main objective of these methods is the analysis of data with respect to the identification of previously unknown and important data patterns. Moreover, the character of these non-trivial methods allows a broad scope of data mining applications and leads to the fact that data mining can be viewed as being detached from specific types of data repositories. This fact is incorporated into a general definition of data mining by Han and Kambler:

“data mining is the process of discovering interesting knowledge from large amounts of data stored either in databases, data warehouses, or other information repositories.” (Han, Kambler 2001, p. 7)

¹ Han and Kambler apply a definition of knowledge which differs from the one used in this thesis.
In addition to this definition it has been shown that data mining distinguishes itself from its predecessors by the type of methods which are available to (automatically) analyze data to identify unknown data patterns which represent interesting information with respect to a specific objective or problem. In order to include these attributes of data mining into a definition, Krahl, Windheuser, and Zick apply a descriptive definition approach:

Data mining is defined as a set of methods with the following performance features:

- Automated forecasting of trends, behavior, and patterns on the basis of known schemes of behavior from the past (controlled learning).
- Automated uncovering of unknown structures from previously unordered data-heaps (uncontrolled learning).
- Additional components supporting the preprocessing (that is the data transfer from a data warehouse, data cleaning, as well as specific analytical functionalities such as the identification of outliers) as well as the visualization and formatting of results.
- Data mining does not contain analytical processes in which SQL queries, report generators, and OLAP tools play a leading part, because they create explicit information. Whereas data mining creates implicit information (Krahl, Windheuser, Zick 1998, p. 24).

4.2.1.2 Methods and Application Scenarios

Within literature the term data mining often comes along with another term, namely knowledge discovery in databases (KDD). Although these terms are sometimes used synonymously, data mining is properly an important stage within the KDD process which might be the reason why these terms are sometimes confused. Figure 4-5 depicts the KDD process and its steps.

![Figure 4-5 The KDD Process (Ester, Sander 2000, p. 2)](image-url)
During the first step of the KDD process, focusing, an understanding for the existing domain knowledge of the respective application area is created. This forms the basis on which the objectives of the application are developed. Additionally, the data which should be mined is defined and, if not already available, collected. This can either be through data selection from an existing database or by the help of explicit data surveys. Historically, the data which should be mined is stored in flat files, but more and more database systems and data warehouses include the necessary functionalities to allow a direct integration in order to facilitate data selection and reduce the effort to sustain consistency.

The preprocessing step deals with the integration, cleaning, and completion of data which is the source of the mining process. In the case that the source data are collected from several databases, an integration of such data is necessary in order to ensure consistency. Reasons for this are, for example, attributes of an object have different names in each source database or values of specific attributes differ among the source databases. Inconsistent attribute values may occur when different departments of a company store values with respect to diverse time frames, e.g. department A uses sales data which is aggregated based on days while department B bases its aggregation on weeks (Ester, Sander 2000, p. 3). An additional task of this step is that incomplete data like empty attributes are specified more precisely (e.g. with the reason why they are empty) or filled up if necessary. The application of data warehouses facilitates the activities of this step, because data are already stored in a consistent and integrated form.

In order to enable the application of data mining methods and algorithms on the selected data, it might be necessary to transform the data to suit the requirements of the methods which ought to be used. This can be achieved either by transformation of the data type, e.g. through discretization (Ester, Sander 2000, p. 4), or by transformation of the values, e.g. performing summary or aggregation operations (Han, Kambler 2001, p. 7). This step is called transformation.

While the previous steps have been preparation, the actual data mining and thus the analysis of the selected data is conducted at this stage of the process. The following most important functionalities of data mining can be distinguished

- Clustering / Cluster Analysis
- Outlier Analysis
- Classification
Chapter 4 – Evaluation of Knowledge Management Technologies

- Prediction
- Association Analysis
- Text Mining

The aim of *clustering* is to divide data objects into clusters (categories, classes, or groups) with respect to the principle of “maximizing the intraclass similarity and minimizing the interclass similarity” (Han, Kambler 2001, p. 25). Obviously, it is essential to model the similarity of the data objects by specifically defining similarities and dissimilarities in order to apply clustering methods successfully. A common way to achieve this definition is through the development of a *dissimilarity distance function* called \( \text{dist} \) (Ester, Sander 2000, p. 46). Such a function is defined for a pair of data objects and results in a metric for the distance between the two objects, where a shorter distance means a higher similarity and vice versa. Usually, attributes of the data objects are used directly or indirectly to define a dissimilarity distance function.

Apparently, the development of an appropriate dissimilarity distance function is essential for the quality of the resulting clusters which may be, at the same time, the most difficult task within a cluster analysis. Therefore, a distance matrix containing all possible data objects is used in some cases in place of a specific dissimilarity distance function.

*Outlier analysis* is the inversion of the clustering analysis. According to Han and Kambler outliers are “data objects that do not comply with the general behavior or model of the data” (Han, Kambler 2001, p. 25). These data objects are detected either by the application of the same function as within clustering or through the use of *deviation functions* which analyze the differences of main characteristics of objects. Although outliers are considered as noise in most data mining functionalities, their identification is essential for applications such as fraud detection.

Contrary to clustering which is applied to identify and create classes of data objects, *classification* uses classes which are defined by the user prior to the execution of any classification methods. It is then the objective of such methods to assign new data objects to the known classes. Before this can be achieved, a given set of data objects, for which the class memberships are already known, is utilized as training data, in other words, the methods analyze the training data with the aim to learn a function which enables the assignment of new data objects to the existing classes based on the data object’s attributes. The resulting function is called *classifier*. Representations of
classifiers can have several forms, e.g. classification rules, decision trees, or mathematical formulae. Figure 4-6 illustrates the classification rules and decision tree representation for an example classifier for risk classification in the domain of car insurance (Ester, Sander 2000, p. 108).

In many cases the classes of data objects are determined by a single attribute which is called the class label attribute. In addition to the development of a classifier through a set of training data, a set of test data is applied to estimate a classifier’s accuracy. Han and Kambler mention that “the holdout method is a simple technique that uses a test set of class-labeled samples. These samples are randomly selected and are independent of the training samples. The accuracy of a model on a given test set is the percentage of test set samples that are correctly classified by the model.” (Han, Kambler 2001, p. 281). Once a satisfying accuracy has been achieved the classifier can be used for the classification of new data objects.

Prediction is a data mining functionality with two different application fields, either the prediction of classes of future data objects or the estimation of missing values or value ranges of data objects. The first application field uses the methods of classification as described above to determine normative values or classes. The prediction of continuous values on the other hand bases on regression techniques.

The last major functionality of data mining is association analysis which is sometimes also referred to as association rule mining. The purpose of this functionality is to identify interesting associations between data objects within large amounts of data, mostly conducted on relational or transactional databases. A famous example for association analysis is market basket analysis which is applied on data of customer transactions in order to identify bundles of products which are often purchased together.
Chapter 4 – Evaluation of Knowledge Management Technologies

(Han, Kambler 2001, p. 226). The resulting associations describe purchasing patterns of customers and can be used to optimize, for example, shelf structures, marketing initiatives, or pricing decisions. An example association rule in this case might be

\[ \{\text{milk, coffee}\} \rightarrow \{\text{cake}\}. \]

The association rule means that customers who buy milk and coffee also buy cake. Obviously, such an association cannot provide one hundred percent accuracy, because it is very likely that some customer who buys milk and coffee leaves the shop without cake. Furthermore, due to the combinatorial way with which associations are built, a number of associations can be generated which is almost as large as the amount of data itself. Thus, two quality criteria are needed in order to identify only those associations which provide interesting information for the decision makers. Krahl, Windheuser, and Zick describe a common way to implement this quality measurement through the usage of the association attributes \textit{support} and \textit{confidence} (Krahl, Windheuser, Zick 1998, p. 81). The two attributes are defined as follows:

\[
support(A \rightarrow B) = \frac{\#\text{transactions containing } A \text{ and } B}{\#\text{transactions in database } D}
\]

\[
confidence(A \rightarrow B) = \frac{\#\text{transactions containing } A \text{ and } B}{\#\text{transactions containing } A}
\]

Thresholds for both of the attributes are defined as the minimum percentage an association has to fulfill, and hence, uninteresting associations can be excluded from the analysis.

\textit{Text mining} is a data mining method which can be used, for example, for keyword-based association and document classification analysis (Han, Kambler 2001, pp. 433-434). Both of these supposedly information oriented analysis procedures, however, can be ascribed to the standard data oriented methods of data mining. In the case of keyword-based association analysis, specific keywords have to be generated for the source documents. The generation of keywords is, yet, closely linked with the domain of information retrieval which will be discussed in section 4.2.3 and is not a component of data mining. The resulting keywords are then stored in the following way (Han, Kambler 2001, p. 433):

\[ \{\text{document_id, a_set_of_keywords}\} \]
After creating such an entry for each source document, the same methods as for market basket analysis are applied to identify keyword associations.

Data mining is followed by *evaluation* which represents the last step of the KDD process. The most difficult activity of this step is the appropriate visualization and presentation of the identified patterns from the preceding step. This depends on the quantity of patterns identified, the pattern type, and the number of attributes of data objects to an enormous extent. In addition to that, the estimation of the validity and significance of the resulting patterns and the derivation of consequences is the other important task which is comprised by this step.

### 4.2.1.3 Analysis and Characterization

As previously explained, the reason for the existence of data mining methods is the increasing amount of available data. Authors in the field of data mining often state that the identification of specific patterns enables the extraction of knowledge which is embedded within databases (e.g. Han, Kambler 2001, p. 4 or Lusti 2002, p. 260). With respect to the definition of knowledge within this thesis it can be argued that the view of these authors is not fully precise. The consideration of data mining applications like market basket analysis, fraud detection, or risk analysis leads to the thought that data mining functionalities enrich data through the identification of patterns or classes in such a way, that a person familiar with the domain is capable of deriving a meaning for the presented results. Hence, domain specific information is generated which then can be combined with other information and knowledge to create new knowledge. But data mining contains no functionality which particularly supports this combination of information and thus the generation of knowledge. This is also the reason why data mining nowadays is mostly integrated into the context of the KDD process as explained above.

On the other hand, approaches to mine more unstructured data as conducted within text mining might cast doubt on the argument above. However, the fact that those methods can be ascribed to the same data oriented methods as used within other data mining applications, as described previously, diminishes any doubt. Considering the continuum of context explication as the dimension for a knowledge management classification, the discussion above can be summarized by assigning data mining to the class *Data Approach*. 
With respect to technology forecasting it could be assumed that, at least in the short run, data mining might be a useful tool to predict future developments. Armstrong, however, argues that “an immense amount of research effort has so far produced little evidence that data-mining models can improve forecasting accuracy.” (Armstrong 2001, p. 10). Thus, the quality of forecasts which solely base on data mining is debatable. Moreover, when taking technology forecasting for strategic innovation management into account, it is questionable whether data mining can be applied at all. It has been explained that data mining basically copes with structured mass-data and additional technologies are needed in order to analyze semi- or unstructured information objects. It is the nature of strategic innovation management, however, that decision-makers face a vast amount of unstructured information which they have to analyze in order to identify trends or developments in the domain of the company. Gaines and Shaw consequently state that “each month the technical press contains a number of major surprises. ... The past is not an adequate determiner of the future in these circumstances.” (Gaines, Shaw 1986, p. 3). This characteristic of technological innovation leads to the result that data mining is neither suitable for the creation of technology forecasts nor for the application within forecasting activities.

However, it is the opinion of the author that data mining can be utilized successfully to facilitate specific tasks within other steps of the technology forecasting process than the implementation of forecasting methods. During the discussion of the technology forecasting process in chapter 3.2 it has been shown that step II and step VIII require the analysis of great amounts of information with respect to specified criteria. In step II information is needed which can be associated with the objectives of the forecast as defined during step I, while an ongoing analysis of information based on the results of a forecast is required within step VIII. In combination with other technologies, data mining might be a suitable way to improve the efficiency of identifying interesting pieces of information through classification and association analysis. Therefore, data mining can be assigned to the classes Step II – Obtain Information and Step VIII – Ongoing Monitoring of the technology forecasting dimension.

4.2.2 Case-Based Reasoning

4.2.2.1 Introduction and Definition of Important Terms
The origins of case-based reasoning (CBR) lie in the field of artificial intelligence. Riesbeck and Schank write, “The search for general mechanisms underlying
intelligence is surely what AI is about and the writing of programs that apply general mechanisms to specific new problems has always been viewed as the ultimate in AI.” (Riesbeck, Schank 1989, p. 1). Case-based reasoning basically has to be seen as a concept which explains the mechanisms and principles of human cognition and reasoning. The characteristics of these mechanisms and principles allow the transfer of case-based reasoning from human cognition to computer systems, for example, in the domain of decision support.

Bergmann et al. define case-based reasoning in a way which is widely accepted in the corresponding literature:

“Case-based reasoning means learning from previous experiences.”

(Bergmann et al. 2003, p. 21)

Previous experiences are essential for the successful solution of new problems and therefore they are the central aspect of case-based reasoning. Riesbeck and Schank analyze human reasoning and argue that “Virtually whenever there is a prior case available to reason from, people will find it and use it as a model for their future decision making. … People reason from experience. They use their own experience if they have a relevant one, or they make use of the experience of others to the extent that they can obtain information about such experience.” (Riesbeck, Schank 1989, pp. 6-7). Examples for this process can be found in many situations in everyday life, e.g. “A financial consultant working on a difficult credit decision task, uses a reminding to a previous case, which involved a company in similar trouble as the current one, to recommend that the loan application should be refused.” (Aamodt, Plaza 1994, p. 40). Besides being based on previous experience, an important aspect of case-base reasoning is the motivation to solve a problem by applying certain experience. Consequently Aamodt and Plaza define case-based reasoning as:

“[solving] a new problem by remembering a previous similar situation and by reusing information and knowledge of that situation.” (Aamodt, Plaza 1994, p. 40)

Experiences (in this context also referred to as solutions) as well as problems are called cases. A case can be anything, e.g. a certain event, an action, a plan, or just a useful piece of information. In the context of human reasoning the access to a previous experience and, hence, to a case is achieved through the activities of remembering and reminding. A computer system needs a variety of functionalities to achieve a similar
way of storing and retrieving cases. Riesbeck and Schank state that “it is the job of the case-based reasoner to have a library of cases; a method of storing new cases that allows them to found again when needed; an indexing scheme that reflects processing that has gone while a case was initially considered; a method of partial matching that allows new cases to be considered in terms of similar ones; and, a method of adaptation that allows information garnered from one case to be applied to another. Human experts have all these abilities hidden under the notion of reminding.” (Riesbeck, Schank 1989, p. 24). The storage from which cases are retrieved is called case base or knowledge base. Obviously, humans possess a very efficient and flexible case base, since they are able to retrieve previous experiences within a fraction of a second independent from whether the case belongs to the same domain as the problem which has to be solved or not. Riesbeck and Schank propose a concept called Memory Organization Package (MOP) which can be viewed as an object oriented data structure and is one example for a dynamic memory model for the implementation of flexible case storage and retrieval (Riesbeck, Schank 1989, pp. 34-36).

It has been shown that case-based reasoning originally can be traced back to the mechanics of the human mind and human reasoning and that the term case-based reasoning today also refers to the artificial implementation of case-based reasoning. Figure 4-7 gives an overview of the general process of case-based reasoning which is also known as the CBR Cycle developed by Aamodt and Plaza (Aamodt, Plaza 1994, p. 45).

---

Figure 4-7 The CBR-Cycle (Aamodt, Plaza 1994, p. 45)
Each iteration of the CBR Cycle starts with a problem, thus with a new case.

**Retrieve** – Case retrieval starts with identifying what is called a *feature*. Features are attributes or descriptors of the new case and are used to “understand” the problem in order to find previous cases with similar features. Therefore, an initial match is made, that is, a set of cases is retrieved within which each case matches the new case to a certain extent. Aamodt and Plaza explain that “cases may be retrieved solely from input features, or also from features inferred from the input. Cases that match all input features are, of course, good candidates for matching, but - depending on the strategy - cases that match a given fraction of the problem features (input or inferred) may also be retrieved.” (Aamodt, Plaza 1994, p. 50). Following this, the most promising cases from the set generated previously are selected. It depends on the application scenario whether it is appropriate to select a single best case or a small subset (Kolodner 1993, p. 19).

**Reuse** – After identification of the single case or the set of cases which are most appropriate for the solution of the problem the selected cases can either be reused directly or, depending on the application and the degree of similarity, an adaptation of the retrieved cases may be needed in order to be able to apply them to the new problem. In some cases the direct reuse of cases is also referred to as *null adaptation* (Riesbeck, Schank 1989, p. 44). Moreover, Riesbeck and Schank distinguish between structural adaptation techniques and derivational ones. Structural adaptation is achieved through the comparison of specific features of cases and the subsequent utilization of a certain set of rules – *adaptation rules* – to transform the solution of the old case towards the new case. The solution for the new case with respect to a derivational technique is created through reusing the same method that has also created the solution for the old case. In other words, structural adaptation techniques operate directly on the old solutions to produce new ones, whereas derivational techniques operate on the method which was used to generate the original solution (Riesbeck, Schank 1989, p. 50).

**Revise** – Case revision comprises the solution evaluation and the repair of possible faults. The evaluation of a suggested solution means nothing more than applying the solution in the real world. Meanwhile the new case remains within the case base with a flag set which identifies it as a non-evaluated case. If the solution was incorrect, errors are detected for which an attempt is made to derive explanations. Therefore, an additional set of domain specific rules, comparable to the ones utilized within adaptation, is used to generate an explanation for the errors found. The explanation is
then applied to guide the repair of the case. However, sometimes errors cannot be identified until the actual problem is solved, e.g. within technical help-desk applications, a component might be identified as being the solution for a problem, but in reality the component works just fine. In such a case, a repair strategy might be to go back within the CBR-Cycle and retrieve a different case (Aamodt, Plaza 1994, 52). In general, failures are stored within the case base in order to prevent these from happening again in future processes.

Retain – Once a solution is confirmed, independent of whether it has been created through a previous case or through external methods (e.g. additional research through an individual), a case is stored within the case base which can be viewed as the learning process of the system. When a solution has been found by the system in the previous steps, the old and the new case are merged into a more general case which incorporates both. If that does not occur, a new case is generated which contains the new solution. Corresponding to the memory structure of the case base, the case which ought to be stored has to be indexed and the case base itself updated. Indexing is an essential problem within case-based reasoning, because it presents the basis of the retrieval process and, hence, the efficiency and applicability of a case-based reasoning system (Kolodner 1993, p. 23).

4.2.2.2 Methods and Application Scenarios

According to Bergmann et al., three general approaches of case-based reasoning applications can be distinguished, namely a textual, a conversational, and a structural CBR approach (Bergmann et al. 2003, p. 21).

Textual CBR targets applications for which documents are essential for the solution of a case. More precisely, each case consists of a document and a short description or summary of that document. The description is necessary to facilitate the retrieval of cases and to keep the number of cases which have been retrieved low. Bergmann suggests that “Otherwise, textual CBR retrieves a large number of cases that are irrelevant.” (Bergmann et al. 2003, p.21). The actual retrieval process utilizes various keyword matching techniques in order to identify similar previous cases, e.g. string comparisons or N-grams. N-grams is a technique which transforms words into lists of N characters which are then used within a matching algorithm. This has the advantage of being robust to mistakes and misspellings which otherwise might lead to a decreased similarity of retrieved cases when applied on long text blocks.
The conversational approach is useful in situations where a problem solution should be achieved by guiding the user. Cases contain questions which then lead to the problem solution. In some applications the integration of these cases into tree-structures is useful in order to group cases and to ask reasonable questions which lead to a solution. Without such a support, a user might be misled due to questions which are inappropriate within the specific context. Bergmann et al. argue that such an approach is useful when simple problems have to be solved over and over again. However, a certain maintenance effort is needed in order to implement this approach, since the case base has to be created and updated manually (Bergmann et al. 2003, p. 23). Typical application scenarios for these types of case-based reasoning systems are call-center and help-desks or online user support systems.

While the conversational approach needs human interaction to integrate new cases into the case base, the third approach – structural CBR – supports the full CBR-Cycle but also comprises the most complex structure of components. General domain knowledge is used to enable matching of features of different cases which are semantically similar (Aamodt, Plaza 1994, p. 49). This structure, also referred to as the domain model, is the underlying basis of the structural CBR approach and specifies the set of features used to describe the different cases. Additionally, such an application allows the use of concepts like MOP, which has been introduced in the previous chapter, to achieve a dynamic case base and to enable a learning process of the system. Thus, a human interaction to update the case base is not necessary. According to Lenz et al., examples for application scenarios of the structural CBR approach are the support of diagnosis like medical diagnosis, of design processes, or of planning activities amongst others (Lenz et al.1998, pp. 12-14).

In general the application of case-based reasoning is useful in situations where the development of a system based on a logical model is not appropriate. Logical models include several drawbacks which make case base reasoners an interesting alternative. Consequently Kolodner argues that “we can’t expect a computer program to be seeded with all the possible combinations of problems it might encounter. Nor can we expect it to have efficient algorithms for generating plausible solutions from scratch all the time.” (Kolodner 1993, p. 24). Moreover, systems which are based on logical models (often called rule-based systems) have trouble dealing with incomplete information. Kolodner writes, “A case-based reasoner makes assumptions to fill in incomplete or missing knowledge based on what its experience tells it, and goes on from there. Solutions
generated this way won’t always be optimal, or even right, but if the reasoner is careful about evaluating proposed answers, the case-based methodology gives it a way to generate answers easily.” (Kolodner 1993, p. 24). It is obvious, however, that the quality of a case-based reasoning system can only be as high as the quality of its case base.

4.2.2.3 Analysis and Characterization

When compared to data mining, case-based reasoning is a concept which targets information rather than data. A case provides the solution to some problem which basically can be viewed as providing domain specific information. The information has been stored within a repository with the intention of being retrieved again when certain features of a new situation, a new case, are similar to the features of the stored case. Up to now, this mechanism is nothing more than a database with structured query methods. However, case-based reasoning comprises certain functionalities which distinguish it from simple database systems and allow the emulation of cognitive processes in order to generate solutions. These functionalities are the capability of adapting old cases to suit the needs of new cases and the fact that a system enlarges its case base by evaluating and retaining cases which have either been solved or provide information about faults. Systems following the structural CBR approach integrate both of these methods and apply general domain knowledge to a model to improve case storage and retrieval, thereby putting the different cases into a certain context. The context is defined by a set of features which are used to index a case and to determine similarity between different cases. Thus, features are descriptors of pieces of information and the corresponding context.

On the other hand, there are also case-based reasoning systems which do not have an underlying domain model, like those which use the textual CBR approach. Such systems work directly on the information and utilize certain algorithms to compare and match new cases with those contained in the case base. Consequently, one has to characterize case-based reasoning with respect to the knowledge management dimension of the Context-Complexity-Matrix as lying between the Descriptor Approach and the Information Approach.

Considering the dimension for technology forecasting, an appropriate characterization and the corresponding identification of the potential for supporting the technology forecasting process is a more difficult task. Reconsidering the argument of Gaines and
Shaw (which has already been stated in the discussion of data mining in chapter 4.2.1.3) it seems that the past is not appropriate for predicting the future in the case of technology and innovations (Gaines, Shaw 1986, p. 3). Case-based reasoning, however, is designed around previous experiences. This leads to the conclusion that case-based reasoning cannot be applied to the execution of technology forecasting activities and thus it cannot be assigned to the class Step IV – Implement Methods – of the technology forecasting dimension. Moreover, taking the requirements of the classes Step II – Obtain Information and Step VIII – Ongoing Monitoring into account, it is doubtful that case-based reasoning is a useful method to support these activities. Both steps have the need to handle a great amount of new information and to put this information into context; either to achieve a clearer perspective of the scope of the forecast or to collect information to monitor the results of the forecast. Case-based reasoning is not a method which is intended for the identification of new information and thus it cannot be assigned to the classes Step II – Obtain Information or Step VIII – Ongoing Monitoring of the technology forecasting dimension.

Nevertheless, it is the purpose of case-based reasoning to support decisions and to solve problems. Due to this it can be imagined that case-based reasoning is applied during Step VI – Prepare Decisions. More precisely, case-based reasoning can be used to support planning activities (Lenz et al. 1998, p. 14). A company which has a long experience in pursuing and developing innovative technologies might profit from its knowledge when a new technology is about to be developed or integrated. Therefore, a case-based reasoner could be implemented which bases on a domain model specific to the domain of the company and contains the plans of how previous innovations and technologies have been developed. Such a system can then be used during decision preparation in order to determine the requirements for the development of a certain technology through the help of previous similar planning situations. Thus, a more complete view of an innovation or a technology and its development can be created which may result in decreased uncertainty. However, such a solution is highly dependent on the definition of similarity, which might be very difficult to achieve and complex within the area of innovation and future technologies.
4.2.3 Information Retrieval

4.2.3.1 Introduction and Definition of Important Terms

The motivation for the development of information retrieval (IR) is straightforward. Increasing numbers of texts, books, or other media led to a greater need for more efficient methods for storing and retrieving this information. This becomes especially obvious when thinking about examples like libraries, archives, or other large information repositories, e.g. the world-wide-web. In all of these cases the amount of information is overwhelming and it can be considered impossible to identify exactly the pieces of information needed by reading all information available (van Rijsbergen 1979, p. 5).

Additionally, for all the mentioned examples their purpose is to satisfy a need for information which most people have in some form (Salton, McGill 1983, p. 2). Reconsidering the fact that the amount of available information increases at an enormous pace, it becomes clear that “essentially, the problem is to find that information in the form of text(s) and other media which optimally satisfies the user’s state of uncertainty and problem space.” (Ingwersen 1992, p. 50). In other words, an individual needs to have control over the retrieval of information in order to retrieve information selectively and thus achieve a higher efficiency and precision in the satisfaction of information needs. Therefore, information retrieval encompasses methods and concepts which support the systematic search and retrieval of information. Moreover, the scope of information retrieval comprises the chain from information organization over storing to retrieval in the end. Consequently, Ingwersen gives the following characterization of information retrieval which will be considered a definition within the context of this thesis:

“Information retrieval is concerned with the process involved in the representation, storage, searching and finding of information which is relevant to a requirement for information desired by a human user.”

(Ingwersen 1992, p. 49)

Naturally, any author of a text wants his work to be as widely available as possible. However, users have to be aware of the existence of information before they can utilize it. In the case of a new book, for example, it does not make sense to send this book to everybody who might have interest in it in order to achieve the necessary awareness. Rather, a representation of the information source is added to some repository which
allows a user to find it when needed. Pao writes, “Information representation is that aspect of information retrieval in which the original file of documents is represented by a set of tags or surrogates such as abstracts or index terms. … The physical forms of representation are organized in such a manner that they may be manipulated and searched to access more efficiently and effectively the content of the collection.” (Pao 1989, p. 98). Thus, representations facilitate organization of an information repository to improve the quality and performance of information search.

The type of representation chosen is closely linked to the search technique of an information retrieval system. Most users search for information by formulating topics or keywords in order to express a certain need for information. Hence, index representations of information can facilitate searching activities by comparing index terms with the formulated topics and keywords of the user. Accordingly Meadow states that “nothing in the entire process is more important to the outcome than the user's own understanding of his or her information need” (Meadow 1992, p. 7).

In addition Ingwersen identifies another conflict with respect to the index and its creation and the perspectives which author and user have regarding a given topic. Ingwersen argues that three different types of aboutness can be distinguished, i.e. author aboutness, indexer aboutness, and user aboutness, where “fundamentally, aboutness refers to the question: ‘What is this document, text or image about?’” (Ingwersen 1992, p. 50). Author aboutness considers the derivation of keywords and terms directly from a document itself which is also the foundation for concepts of, for example, automatic indexing and text analysis. Indexer aboutness approaches the question stated above from a point of view which focuses the representation of information. Ingwersen write, “In this case, the question of what a document is about may be answered by an indexer who, by means of classification or indexing of documents, attempts to summarize or surrogate the contents of the message in each document or piece of text.” (Ingwersen 1992, p. 51). The meaning of a given piece of information and the indexer’s general knowledge of the corresponding domain influence the derivation of index terms or other surrogates. Additionally, in many cases it is the objective of indexing activities to unify the interpretations of representations with respect to a certain domain. Therefore, the resulting terms summarizing the meaning of a piece of information which have been created from an indexer’s perspective might diverge from the original intentions of the author. User aboutness, on the other hand, is determined by a user’s context which is variable and often deficient. Ingwersen argues that “the user can express what he knows,
but cannot define the state of uncertainty and problem space in terms of information, the nature of which the user does not yet know.” (Ingwersen 1992, p. 52). Therefore, user aboutness also differs from the two other aboutness perspectives.

It is essential to take these varying perspectives of aboutness into consideration when developing an information retrieval system in order to be successful. Consequently, an information retrieval system requires certain intermediary functionalities which support a user in defining search requests by understanding the user’s problem space and its domain before transforming these requests into queries to the information repository, while queries can be understood as “formal statements of information needs” (Frakes, Baeza-Yates 1992, p. 1). This transformation of user requests for certain information into queries is executed with respect to the logic used for indexing and representation of pieces of information, hence, connecting user and information repository and thereby facilitating the searching and finding processes. Figure 4-7 summarizes the model of information retrieval as described in the discussion above.

4.2.3.2 Methods and Application Scenarios

Application scenarios for information retrieval systems are manifold. Some examples like supporting the management and retrieval of information within libraries and archives have already been mentioned. Frakes and Baeza-Yates state that “any discipline that relies on documents to do its work could potentially use and benefit from IR.” (Frakes, Baeza-Yates 1992, p. 1). Hence, obvious additional application fields might be support functionalities within research facilities, public institutions, or certain departments of companies, since today most forms of intellectual and creative work include the internalization and processing of information in the form of texts, documents, pictures, or other media. In the case of research facilities, an information...
retrieval system for research articles, patent information, opinion papers, test results, or other reports might be capable of improving the quality of the researchers’ work by supporting the identification and retrieval of necessary information when needed. Advertising agencies might benefit from an information retrieval system which focuses on images and other advertising related material and thus, reduces the time necessary to identify and retrieve needed material. In general, the idea of applying information retrieval systems is to support users to find specific information in a large amount of available information in order to satisfy their personal information need. Consequently, due to the trends in today’s information society and the still increasing importance of the world-wide-web, the number of possible application areas for information retrieval will increase further.

On the theory side, Frakes and Baeza-Yates argue that literature came up with several different taxonomies for information retrieval conceptual models. However, “the problem with these taxonomies is that the categories are not mutually exclusive, and a single system may contain aspects of many of them.” (Frakes, Baeza-Yates 1992, p. 3) General concepts for information retrieval systems include text-pattern search, signature search, and Boolean search which is also referred to as inverted file search.

The implementation of an information retrieval system with text-pattern search as its foundation is only reasonable for a small number of texts, because of the necessity to access each text successively in order to analyze it with respect to a certain text-pattern. The input text-pattern is defined as the query and can either be a text string or a regular expression and thus, allowing a user to conduct precise searches for a certain text-pattern as well as inexact searches through the utilization of wildcards or other operators within regular expressions, i.e. using the regular expression “Uni*”, the words “University”, “Unit”, and “United-Nations” are examples for possible matches. Algorithms for text-pattern search apply this comparison scheme to identify texts which match the query of a user. Therefore, text-pattern search can be considered as one form of full-text search. However, such an approach toward searching can be very time consuming in the case of large repositories or documents.

Signature search bases on the generation of files containing a signature for each text document within the repository. A given document is divided into logical blocks, while each logical block comprises a constant number of distinct, noncommon words (Frakes, Baeza-Yates 1992, p. 46). Words like “the”, “a”, or “from” are common among
documents and, hence, do not have any distinctive value. In order to improve precision, these words are left unconsidered with respect to the resulting signature. For each noncommon word of a logical block a bit pattern of a fixed length is generated. Such word signatures are then combined using an OR operator for each logical block to form block signatures. The concatenation of the single block signatures results in the document signature. Search requests for specific words are executed through the creation of a signature for the query word and its comparison with the different block signatures of each document. The word is contained within a logical block, when for each position with a “1” in the word signature there is a “1” in the corresponding position of the block signature. Apparently, the precision of such an approach depends on the quality of the generation of word signatures, because a weak function might lead to words being identified in a logical block, although they are not contained. The reason for this is that a block signature might have “1”s in the corresponding positions due to OR combinations of other words.

The most widespread concept of information retrieval is Boolean or inverted file search. Both terms most often refer to the same concept which applies a keyword representation of documents in order to cope with large amounts of information. Keywords describing the content of a document are created and then organized in a file in an inverted fashion, that is, keywords point to references of documents in which they appear (Pao 1989, p. 156). Searching is implemented by looking-up query terms in the inverted file without any access to the actual documents needed. Thus, such a representation allows much more efficient search functionalities and, thereby, is also applicable to large information repositories. Inverted files are one form of index organization. The Boolean aspect of this information retrieval concept regards the formulation of queries. Keywords in a query can be connected through the Boolean logical operators AND, OR, and NOT in order to specify or generalize a search request (Frakes, Baeza-Yates 1992, p. 3).

Clearly, the creation of an index whether it is organized in an inverted file or uses another file structure requires a certain effort. According to Pao, indexing is a process which comprises two separate steps, the analysis of the text which needs to be indexed and the selection of index terms which describe the respective document (Pao 1989, p. 131). In case of a large information repository, these activities can be very time consuming. This led to the development of methods and algorithms for automatic indexing or automatic text analysis. Van Rijsbergen distinguishes between two general approaches of automatic text analysis, namely, statistical and linguistic (semantic and
syntactic) approaches (van Rijsbergen 1979, p. 14). The breadth of the evolved methods ranges from simple indexing which is based on content words in a title, up to complex probabilistic text analysis. A simple example for a statistical indexing method is based on the number of a single word’s occurrences which is determined by counting each occurrence. The words and their corresponding number of occurrences are organized in a list in descending order of occurrences. Additionally the count of each word is transformed into a percentage value. These percentages are then compared to the statistical average occurrence percentage of a word within a text. In the case that the percentage of the word is higher than the average, the word is considered important for the text and, therefore, chosen as an index term.

Another way of indexing information which uses a completely different approach is citation indexing. This approach follows the idea that references within a text to another text express a semantic relation between both texts. Furthermore, the higher the reference frequency of a text is, the more important is the work, while reference frequency can be understood as the number of texts referring to the work (Pao 1989, p. 132). On the other hand, it cannot be assumed that all citations and references have the same relevance and weight. Pao, for example, argues that “citing theoretical work to support an experimental design is substantially different from citing a reference of historical interest to the research problem.” (Pao 1989, p. 132). Therefore, differentiated analysis of citations and references within a document might lead to a more precise statistical weight of references and thereby improve the importance ranking of documents and the quality of the information retrieval system. Nevertheless, the fact that such an approach adds practical value to an information retrieval system can be observed by examining online services like Google’s Scholar or the digital library of the Association for Computing Machinery (ACM). Both services utilize citation analysis and indexing in order to provide more sophisticated search functionalities of scientific papers.

Besides the efforts spend in automatic indexing, the last couple of years have shown alternative developments with the focus on the information side. The idea is to enrich a piece of information with explicit contextual information by the help of descriptors in order to facilitate context related searches. Examples for such approaches are the Resource Description Framework (RDF) or the Dublin Core metadata element set.
4.2.3.3 Analysis and Characterization

During the development of the characterization scheme for knowledge management technologies the continuum of context explication has been introduced as a means to distinguish technologies with respect to whether they target data or information. Even more importantly, the continuum facilitates determining to what extent knowledge management technologies make use of contextual information to provide methods for information search and discovery. Considering the definitions for each single class of the knowledge management dimension, it seems obvious that information retrieval belongs to the class Information Approach. In general, such a classification appears to be reasonable as information retrieval targets raw information. Smolnik, Kremer, and Kolbe argue that although information itself comprises content and context, the context is interwoven with the content and thus difficult to explicate (Smolnik, Kremer, Kolbe 2005, p. 37). As a result, technologies which do not include additional explicit contextual information rely only on the content or its representation for the use within search functionalities. Clearly, this is true for a bigger part of the existing information retrieval conceptual models. Moreover examples like signature search transform pieces of information into data and then utilize pattern identification methods, rather than applying methods directly on the information itself. However, the fact that such a method also includes a user defined query distinguishes it from the class Data Approach, because queries describe the content desired by a user and thus, are implicitly related to some form of context.

On the other hand, one can argue that some forms of information retrieval also integrate explicit contextual information into search and retrieval methods. While Smolnik, Kremer, and Kolbe state that “authors have to provide [explicit contextual] information at the time of creation” (Smolnik, Kremer, Kolbe 2005, p. 37), the previous discussion of aboutness allows an additional perspective. It has been shown that three different types of aboutness can be distinguished, i.e. author aboutness, indexer aboutness, and user aboutness. While author aboutness can be viewed as resulting from an author’s context and therefore as being expressed by the author’s creation of explicit contextual information, queries are the outcome of a user’s effort to describe his assumed aboutness of information which satisfies the existing information need in a perfect manner and, hence, result from a user’s specific context. Since such definition of explicit contextual information through authors themselves is not integrated into information retrieval models and because of a user’s context being only implicitly
contained within query formulation, search and retrieval methods do not apply explicit contextual information with respect to author and user aboutness. However, when considering the fact that some information retrieval systems are based on the creation of index terms through analysis of documents and alignment with a specific domain by individuals, one can argue that such indexes represent the indexer’s aboutness and thereby also the context of the individual who analyzes the documents and creates the index. Nevertheless, in the same way that indexer aboutness differs from author aboutness, the contexts of author and indexer vary. Clearly, the often mentioned myriads of available information make such an approach to indexing and information retrieval efficient in only a small number of cases.

In general the characterization of information retrieval by assigning it to the class Information Approach within the knowledge management dimension is a reasonable outcome; however the exceptions as discussed above should be taken into account. Therefore, information retrieval will be categorized as lying in between of Information Approach and Descriptor Approach with the bigger part lying on the Information Approach side.

The discussion of information retrieval up to this point should have given the reader an understanding of when it is useful to utilize information retrieval systems and how information retrieval can be characterized with respect to knowledge management approaches. Within technology forecasting it has been shown that certain process step classes include the need to identify information within a large amount of it being available, namely, Step II - Obtain Information, Step IV – Implement Methods, and Step VIII – Ongoing Monitoring. The difference between the information need of these steps is that the first two steps require a broad range of new information with respect to the selected forecasting scope, while the latter step utilizes specific information which is closely linked with the developed technology forecasts in order to compare them to reality. Therefore, an efficient way to identify and assess relations and consequences of pieces of specific information is more important than just the identification and retrieval of interesting information from a large amount of various, available information. It is a common assumption among researchers of information retrieval that searching within such systems is an iterative process. A user starts with some sort of query and evaluates his own understanding of the information needed with the help of the first result set. Either the information is sufficient – it results in the retrieval of additional information through references or alike – or a user realizes that the request has to be revised over all
Chapter 4 – Evaluation of Knowledge Management Technologies

(Salton, McGill 1983, p. 3). Reasons for this can be found when taking into account that users are only able to describe what they need based on what they already know. These arguments lead to the conclusion that information retrieval is not applicable to Step VIII – Ongoing Monitoring of the technology forecasting process and, instead, can be characterized as being able to support steps with a need for a wide range of new information, thus Step II – Obtain Information and Step IV – Implement Methods.

4.2.4 Topic Maps

4.2.4.1 Introduction and Definition of Important Terms

When it is appropriate to refer to data mining as a natural step in the evolution of information and database technology, topic maps can be viewed as a similar development with respect to information indexing, search, and navigation. The need to merge different indexes was one of the main reasons for the initial development of the topic map concept in 1991 (Rath, Pepper 1999, p. 3). In addition to that it has been stated that the most important fact in this context is that “indexes, if they have any self-consistency at all, conform to models of the structure of the knowledge available in the materials that they index. But the models are implicit, and they are nowhere to be found! If such models could be captured formally, then they could guide and greatly facilitate the process of merging modeled indexes together.” (Rath, Pepper 1999, p. 3). This reveals the motivation behind the development of topic maps; the creation of a technology to explicitly use the “structure of the knowledge” which is implicitly contained in indexes like back-of-the-book indexes. Once available, merging indexes as well as finding information and related objects can be facilitated for which the concept of topic maps is a proposed solution. The various components that comprise a topic map will be defined and explained in the remainder of this chapter.

The most important components of a topic map are topics. The official definition for a topic is as follows:

“An aggregate of topic characteristics, including zero or more names, occurrences, and roles played in associations with other topics, whose organizing principle is a single subject.” (ISO/IEC 13250:2002, p. 5)

Although the definition above describes the characteristics of a topic precisely, it is quite difficult to imagine what a topic stands for. Pepper gives a more general description of what a topic is.
“A topic, in its most generic sense, can be any ‘thing’ whatsoever — a person, an entity, a concept, really anything — regardless of whether it exists or has any other specific characteristics, about which anything whatsoever may be asserted by any means whatsoever.” (Pepper 2002, p. 6)

Pepper adds that the explanation above is closely aligned to the definition of the term subject in the ISO/IEC 13250:2002 standard. Basically, a topic is the virtual representation of a subject within the topic map, while a subject can be considered as any “thing” in the real world. Examples for topics can be “University of Paderborn”, “Paderborn”, or “Business Computing”. As included in the official definition, topics have three characteristics: names (sometimes also referred to as base names), occurrences, and roles played in associations with other topics. Names are terms describing a topic’s subject in human readable form. In most cases, the name of the subject is chosen as such a term. It is not mandatory for a topic to have a name. It is more important to point out that also several names can be assigned to one topic with the constraint that all names refer to the same subject. This mechanism facilitates integration of synonyms or different languages into a single topic.

Occurrences of a topic are pointers to specific information resources which are relevant for the topic. A single occurrence can either be a connection to a resource or a reference to a string value. While the first method is used to bind certain information resources to a topic, the latter can be utilized to assign string values to a topic which, for example, is applicable to enrich a topic with additional meta-data (Rath 2003, p. 13). Additionally, occurrences are categorized through their association to occurrence roles, e.g. “article”, “illustration”, or “definition”. In case an information resource fits into several occurrence roles, an occurrence is assigned to the corresponding topic for each of these roles. The standard remarks that occurrences can also refer to offline resources instead of just pointing to those which are online (ISO/IEC 13250:2002, p. 15). Figure 4-9 illustrates the concept of topic maps up to this point.
Besides the vertical connections between topics and their assigned resources through occurrences, associations link topics horizontally and thus create a network which describes the relations between the topics available. The idea behind the introduction of associations is explained by Rath, who states that “associations simulate the way humans think and as such are essential for knowledge modeling” (Rath 2003, p. 14). The number of topics which are connected through one association is not limited, although most often two topics are associated. Comparable to occurrences, also associations can be categorized by their type. Examples for association types are “is in” or “is part of” which can result into the following example associations when applied to the aforementioned examples for topics:

- University of Paderborn is in Paderborn
- Business Computing is part of University of Paderborn

In general associations are valid independent of the directions they are traversed. This fact might lead to the problem that the roles of each topic involved become unclear. In the example “University of Paderborn is in Paderborn” it is quite clear that the topic “University of Paderborn” plays the role university, respectively the topic “Paderborn” the role city. However, in other cases the assignment of roles can be more difficult. Therefore, the concept of topic maps includes so-called association roles as the third characteristic of topics, in order to assign a specific role to each topic of an association and, hence, facilitate the understanding and interpretation of an association. Figure 4-10 integrates associations into the illustration of the topic map concept.
Besides the characteristics of a topic as explained so far, a topic can be categorized by the *topic type* which it belongs to. A conceivable topic type for the topic “Bach”, for example, is “composer”. However, which topics and topic types are chosen for certain topic map depend on its application. A thesaurus clearly needs different topic types than an application for legal publishing (Pepper 2002, p. 7). In general topic types are also defined as topics within a topic map.

The basic structure of the topic maps as presented above can be enriched by the integration of the additional concepts of *facet* and *scope*. The term *facet* refers to metadata which is assigned to information resources in order to provide more descriptive information about each information resource. More important, however, is the concept of scope. A topic’s characteristics name, occurrences, and associations are extended through the assignment to a certain set of topics, called a *theme*. This is used to define in which context the characteristic is valid and thus, different views of a single topic map can be achieved. A common example for the usefulness of scope is within multilanguage applications. Scope attributes assigned to names and occurrences of topics which define the language of the corresponding characteristic enable a dynamic switch of the whole topic map between the provided languages. Rath writes, “Scopes help to filter the ‘noise’ in large topic maps and allow us to concentrate on the interesting parts. They help to build semantic slices through the topic map.” (Rath 2003, p. 18).
4.2.4.2 Methods and Application Scenarios

Although some people consider the concept of topic maps as being “still well ahead of its time” (Kay 2004, p. 26), several methods already exist for the creation of topic maps, their representation through data structures, and their integration into applications.

According to Garshol three main approaches can be distinguished for the creation of topic maps:

- “Have humans author the topic maps manually. This usually gives very high-quality and rich topic maps, but at the cost of human labor. This is appropriate for some projects, while prohibitively expensive for others.”

- “Automatically generate the topic map from existing source data. This can give very good results if the existing data are well-structured (sound familiar?); if not, there are various natural-language processing tools that might help.”

- “Automatically produce the topic map from structured source data like XML, RDBMS, LDAP servers, and more specialized applications.” (Garshol 2002)

Hence, the creation of a topic map can either be achieved manually or automatically, while the methods used for automatic creation depend on the type of information resources available, with both approaches having specific advantages and fall-backs. The quote above emphasizes that topic maps which have been created manually require a certain amount of human labor cost. This is not only true for the creation of a topic map. Rather, one can expect additional cost during the time the topic map is in active use due to maintenance reasons and needs to keep the system up-to-date. Since information most often depends on time and context to be valuable, only a few types of information systems do not have to cope with dynamic and frequent change, which means that in most cases certain effort has to be spent in an ongoing fashion in order to maximize the value of a certain topic map. Automatic approaches to generate a topic map are one means to decrease such maintenance cost and to achieve a more efficient and dynamic process of generating and maintaining a topic map. Actual methods which are applied to achieve an automatic extraction of topics and the generation of a topic map are comparable to those used in automatic indexing as mentioned in chapter 4.2.3.2. However, methods for automated topic map generation integrate mechanisms to extract
semantic relations in addition to certain topics. Folch and Habert, for example, describe an approach through so-called \textit{Inductive Semantic Acquisition Methods} (Folch, Habert 2000). This approach is based on statistical data derived from multiple heterogeneous and idiosyncratic information repositories in order to identify semantic classes which are used to generate a topic map. Other approaches include the utilization of natural-language processing tools as pointed out by Garshol (Garshol 2002).

The representation of a topic map within a computer system depends on the type of storage system used to some extent. A common data-structure for the digital representation of topic maps exploits the Extended Markup Language (XML) and is defined in the TopicMaps.Org specification \textit{XML Topic Maps (XTM) 1.0}. Through this method, topic maps can be stored as structured text within plain files as well as being imported into database systems for more efficient access. For a complete description of the XTM framework the user is referred to the corresponding part of the XTM specification.

Obviously, some effort is needed to integrate an XTM topic map into an application which supports more functionality than only reading the topic map. In order to facilitate such undertakings, Application Programming Interfaces (APIs) as known from other software development fields are needed. Moreover, through a common standard API the interchange of information and of complete topic maps can be supported and thus, an easier development of topic map applications achieved. A project which tries to achieve the creation of a standard API specification for the development of topic map applications is TMAPI. The project has the motivation to “do for topic maps what SAX and DOM did for XML - provide a single common API which all developers can code to and which means that their applications can be moved from one underlying platform to another with minimum fuss.” (TMAPI 1.0 SP1).

One can see that several efforts are made to achieve a common base for the development of topic maps, because the application scenarios for topic maps are diverse. One application field for manually created topic maps is document categorization and classification. Indexing methods are used to identify whether a document can be assigned to one or more topics of the existing topic map. The result is a facilitated navigation of the underlying information repository and hence, the improvement of information discovery. Topic maps are also useful to provide a semantic approach to search methods instead of simple index based methods. Furthermore, topic maps can be
applied to visualize and enable the navigation of knowledge structures incorporated in so-called *Ontologies* which will be discussed in chapter 4.2.5.

### 4.2.4.3 Analysis and Characterization

Charles F. Goldfarb, the inventor of the Standard Generalized Markup Language (SGML), once said about topic maps that they are “the GPS of the information universe” (Rath 2003, p. 8). This image alludes to the functionality of the General Positioning System (GPS) to locate one’s current position in order to facilitate orientation. Although the comparison of topic maps with GPS seems to be quite reasonable, topic maps are more than just a means to identify a position within an information space. Rather, to stay in that image, topic maps combine the GPS functionality of providing orientation with a grid which lies on top of the environment surrounding oneself, thus, allowing immediate perception of which possible paths exist and where they lead. Returning to the field of information and their repositories, topic maps provide methods to navigate associatively over large amounts of available information in a conscious manner and thus, enable a systematic discovery of information and acquisition of new knowledge by the user. This is possible through detaching the information source from the context used to find the information which results in topic maps being “information assets in their own right, irrespective of whether they are actually connected to any information resources or not” (Rath, Pepper 1999, p. 9). Therefore, topic maps can be viewed as encompassing and merging different information repositories underneath one explicit context, independent of whether the context is created manually or through automatic procedures. Moreover, topic maps support “managing the meaning of the information, rather than just the information” (Garshol 2002, p. 2).

It is clear that an explicit context, a so-called *meta-context*, is used to organize available information in such a manner that more efficient search methods can be applied. Rath, for example, calls topic maps “intelligent search indexes” and argues that they provide better search results than search methods which are based on full-text indexes (Rath 2003, p. 7). Hence, the meta-context is the most characterizing aspect when discussing topic-maps and thus, topic maps clearly belong to the class *Meta-context Approach* when considering the knowledge management dimension of the Context-Complexity-Matrix.
The paragraph above gave an idea about the benefits of topic maps and how they can be characterized considering the knowledge management dimension. It has been shown that topic maps provide methods to deal with the challenge of increasing amounts of available information by creating an explicit meta-context which encompasses a variable amount of information repositories. Clearly, such an approach requires greater efforts in terms of topic map creation and maintenance than simple technologies like full-text search. Although topic map creation and maintenance can be partially automated, certain effort is still needed in order to evaluate the generated map and to customize it to specific needs. When thinking about the technology forecasting dimension of the Context-Complexity-Matrix and the technology forecasting process itself, the problem of coping with large amounts of information and, especially, the identification of valuable pieces of information within these amounts, has to be considered a central aspect with the potential of increasing the quality and efficiency of the technology forecasting process. *Step II – Obtain Information* of the process, for example, has been characterized by the mentioned need to manage such large amounts of information. The challenge within this step is to identify those pieces of information which belong to the scope of the forecast and are therefore valuable for later process steps.

Reconsidering the fact that a topic map stands for a certain context, a topic map can be very useful when created in order to represent the scope of the forecast. Such a topic map comprises the different technologies and research areas within the focus of the company which conducts the forecast. Associations can be used to link technologies in order to express influences and relations among those technologies. Any information to which the topic map is applied to can then be categorized with respect to the scope of the forecast and thus, the identification of valuable information is facilitated. Furthermore, once a comprehensive information repository exists, the topic map can be used to relocate information and to relate it with results of the forecasting activities. Hence, topic maps also provide additional value when used within *Step IV – Implement Methods* and *Step VI – Prepare Decisions*. To be able to identify specific information which correlates with the results of the forecasting activities is especially important within *Step VIII – Ongoing Monitoring*, because each influence on the correctness of the forecast can have enormous consequences for a company’s competitive advantage and thus, its success. Therefore, minimization of the time span between the occurrence of an event which influences the correctness of a technology forecast, like the publishing of
Chapter 4 – Evaluation of Knowledge Management Technologies

information about some disruptive technology, and its identification by the company is essential. The aforementioned filtering and localization capabilities of topic maps help to achieve a more precise analysis of available information and hence, a more efficient monitoring process overall. In general, topic maps have the potential to increase the efficiency of each technology forecasting step in the Context-Complexity Matrix, because they can be tailored to the scope of a forecast and thereby reduce the complexity of available information to a manageable level.

4.2.5 Ontologies

4.2.5.1 Introduction and Definition of Important Terms

The origins of the term ontology can be found in the ancient Greek philosophy. Since then ontology refers to the branch of philosophy concerned with studying the nature of being and belongs to the field of Metaphysics. Research in the area of artificial intelligence led to the need of a formalized way to describe the perception or the view of the world an artificial intelligence system, also called agent, incorporates. Furthermore, a common methodology to express an understanding of a certain domain in a formalized manner was essential to facilitate interaction among these systems through sharing a common understanding as a basis for communication. The term ontology was adopted to represent this formalized structure, describing terms, relations, and axioms which create some sort of artificial perception and understanding. While philosophers always refer to ontology as one singular phenomenon, an arbitrary number of ontologies can be created when used as a means to achieve common understanding and enable communication among different agents or individuals within various knowledge domains. This leads to the following definition of ontologies:

“An ontology is a shared and common understanding of some domain that can be communicated across people and computers” (Benjamins, Fensel, Gómez Pérez 1998, p. 2)

The definition raises the question of how such a shared and common understanding can be implemented. The answer for that question is conceptualization.

“A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose.” (Gruber 1993, p. 199)

Consequently, a conceptualization can be viewed as a model to a certain extent as both focus on selected aspects of reality. It comprises objects and concepts of a specific
domain as well as relationships between these entities. For that reason, Gruber proposes the following definition of ontology:

"An ontology is an explicit specification of a conceptualization." (Gruber 1993, p. 199)

Zelewski, Schütte, and Siedentopf extend the statement of the definition by discussing available methods with which an explicit specification as mentioned in the definition of Gruber can be achieved. The authors argue that the definition lacks a particularization of which specification methods can be used. They stress that besides terminological and syntactical methods, semantic specification methods are predominantly used to create ontologies and hence, are an important aspect when characterizing ontologies (Zelewski, Schütte, Siedentopf 2001, p. 191).

In general, ontologies can be viewed as providing a model or a structure for some specific (domain) knowledge. The center of this model is formed by a set of terms – the terminology of the considered knowledge. Relations between those terms and axioms for derivations among them extend the terminology and are a means to allow automated computational inference. All mentioned components are organized within a taxonomy and build the structure of every ontology. The terms of an ontology’s vocabulary are referred to as concepts and can be characterized through their position within an ontology’s hierarchy. Hovy describes these characteristics in the following way: “We define an ontology simply as a taxonomized set of terms, ranging from very general terms at the top (allowing nonexpert users to find access points) down to very specialized ones at the bottom (allowing them to be connected to specific columns in databases).” (Hovy 2003, p. 48). Additionally, concepts of an ontology are characterized by a set of attributes, while the type of those attributes as well as their values depend on each single concept. Comparably to object oriented programming languages, concepts inherit the attributes of parent concepts which lie above them within the hierarchy.

According to Zelewski, Schütte, and Siedentopf, several other types of ontologies, besides those representing domain knowledge, exist; namely, common-sense ontologies, representation or meta ontologies, task ontologies, and method ontologies (Zelewski, Schütte, Siedentopf 2001, p. 195).

**Common-sense ontologies** – General views of the world which are independent of specific domains and commonly considered as self-evident prerequisites for basic actions and argumentations are incorporated within common-sense ontologies.
Although difficult to generate, such ontologies are of high interest in the field of artificial intelligence.

*Representation / meta ontologies* – Such ontologies are utilized to determine the possible expressions of representation or modeling languages. One example is the so-called “Frame-Ontology”.

*Task ontologies* – Tasks of a general, domain overarching nature with the attribute of occurring as well as being conducted always in the same fashion, independent of the application domain, are specified with the help of task ontologies.

*Method ontologies* – Method ontologies provide the terminology and a corresponding syntax with the help of which those problem types can be defined for which a certain solution method is applicable.

Since ontologies are used to enable communication between any combination of various computer systems, software agents, and individuals, researchers agree that an ontology has to be created in a collaborative manner to ensure its acceptance among all parties involved, e.g. developers, users, or managers. Hence, the applicability and value of an ontology essentially depends on what is called *ontological commitment*, the acceptance of the correctness of the ontology from individuals within the respective domain (Hesse 2002, p.478). Furthermore, ontology creation or design is a sub-activity of *ontological engineering* which “has as its goal effective support of ontology development and use throughout its life cycle – design, evaluation, validation, maintenance, deployment, mapping, integration, sharing and reuse.” (Gruninger, Lee 2002, p. 40).

### 4.2.5.2 Methods and Application Scenarios

As described previously, the interest for ontologies outside of artificial intelligence research rose with the emerging awareness that ontologies facilitate communication by formally and unambiguously specifying those parts of domain knowledge, which are essential for mutual understanding. Situations which benefit from a specified base for communication and understanding are manifold and can be found in many areas of business life. Considering a company’s organization one can most often find examples with two departments using their own terminology in which identical terms have differing semantic meanings or identical semantic meanings are referred to by differing terms. Communication among those departments can be facilitated through the integration of an ontology which can either be used to translate terms into those
corresponding to the targeted department or to merge both terminologies and unify their usage.

Multi-agent-systems are another field which is driving the development and application of ontologies forward. Naturally, a distributed system which encompasses an arbitrary number of autonomously acting software agents needs some form of standardized view of the world in order to enable the interaction and communication of such agents. The coordination of production and logistics processes which involves several firms or the creation of an automated electronic market have an enormous potential for the application of ontologies, since knowledge and its structures vary even more between different companies than between departments of a single firm. Based on examples like those mentioned above, Zelewski derives a general thesis about which aspects lead to a situation benefiting from the application of ontologies:

The stronger the accomplishment of collaborative company tasks

a. bases on knowledge intensive processes and

b. depends on the interaction of actors with at least partially divergent knowledge backgrounds from inside as well as outside the company

the higher the significance of ontologies as a means for integrating task relevant knowledge components. (Zelewski 2001, p. 4)

All of the examples mentioned so far are related to the fact that ontologies are a way to facilitate communication. However, such situations do not always use the full potential of ontologies and, obviously, additional application scenarios exist which also benefit from ontologies. Gruninger and Lee distinguish three general uses of ontologies with communication being one of them:

- **communication**: between implemented computational systems, humans, as well as between implemented computational systems and humans

- **computational inference**: for internally representing and manipulating plans and planning information and analyzing the internal structures, algorithms, inputs, and outputs of implemented systems in theoretical and conceptual terms
reuse and organization of knowledge: for structuring or organizing libraries or repositories of plans and planning and domain information (Gruninger, Lee 2002, p. 40)

An application scenario which received a lot of attention during the last couple of years is called semantic web. Termed by Tim Berners-Lee, the inventor of the World-Wide-Web (WWW), the semantic web describes an extension of the WWW towards a more sophisticated network which supports human information discovery as well as automated, semantic transactions through distributed, autonomous software agents. In fact, to enable automated information processing is the motivation for the development of a semantic web (Berners-Lee, Hendler, Lassila 2001). In order to achieve this functionality of the web, descriptor languages like the earlier mentioned Resource Description Framework are utilized to annotate web pages to enrich them with explicit contextual information. This information can then be used by software agents to analyze and process the information on a web page. Ontologies are used in this context to provide the necessary information structure which allows software agents to “understand” the information contained in the page annotation through direct look-ups within an ontology’s taxonomy or by the help of inference rules which are integrated in the ontology (Berners-Lee, Hendler, Lassila 2001).

But how are ontologies like those described above implemented? One of the latest developments in this field is the OWL Web Ontology Language which is developed and recommended by the World Wide Web Consortium (W3C) and is an advancement of the ontology language DAML+OIL. DAML+OIL is an extension to RDF and evolved from the merge of DAML, the DARPA (Defense Advanced Research Projects Agency) Agent Markup Language, with the Ontology Inference Layer (OIL) (McGuiness et al. 2002, p. 72). The first version of DAML+OIL was released in December 2000 and was developed to provide an XML based method for ontology creation and use within software agents. In the context of semantic web research the need for a language for web ontologies emerged. Existing methods did not suffice for the satisfaction of this need which led to the development of OWL. “If machines are expected to perform useful reasoning tasks on these documents, the language must go beyond the basic semantics of RDF Schema.” (OWL 2004). However, like RDF, OWL is an annotation language for web pages which applies XML tags to enrich documents with explicit contextual information describing meaning and semantic relations and allowing automated information processing as well as more sophisticated search methods through
precise context specifications and computational inference. Figure 4-12 depicts a simple, fictitious example for web page annotation with references to an ontology.

![Figure 4-11 Web Page Annotation for Ontology References](image)

OWL comprises three sublanguages which help with the annotation of web pages. Each of these sublanguages integrates an increasing range of features: **OWL Lite, OWL DL,** and **OWL Full.** While OWL Lite supports the generation of classification hierarchies and simple constraints, OWL DL allows the usage of all OWL language constructs, but includes certain restrictions for this usage in order to retain computational completeness. OWL DL is named in correspondence to the research field of **Description Logics** which studies the logics which are the formal foundation of OWL. OWL Full revokes all restrictions and provides a feature range and leaves the consistency validation to the user (OWL 2004). For more information on the functionalities of the OWL Web Ontology Language the reader is referred to the corresponding W3C recommendation.

### 4.2.5.3 Analysis and Characterization

The previous two chapters gave an introduction to ontologies, their capabilities, and possible application scenarios. It should be clear that ontologies are a means to provide a resource which unambiguously determines the meaning of terms and their relations to other terms within a certain domain. This structure is an autonomous construct without links to specific information resources which has the purpose to make ontologies mobile and reusable. However, it has also been shown that an ontology can only be successful when it is developed in a collaborative fashion and broad commitment within the target domain can be achieved. Therefore, ontology creation, especially with respect to
business applications, requires a certain effort for coordination and cooperation between all parties which are involved and cannot be achieved automatically. This problem of ontology application will exist as long as standard ontologies for specific domains are not developed. In fact, it is a difficult task to create such standard ontologies. Gruninger and Lee give a reason for the difficulties and the restricted reusability which exist when creating an ontology. The authors state that “Although ontologies were originally motivated by the need for sharable and reusable knowledge bases, the reuse and sharing of ontologies themselves is still limited because the ontology users (and other designers) do not always share the same assumptions as the original designers.” (Gruninger, Lee 2002, p. 41).

With respect to the knowledge management dimension of the Context-Complexity-Matrix, it is quite obvious that ontologies belong to the category Meta-context Approach. The reason for this is that explicit context structures are created which are, as already mentioned above, independent of specific information resources and can be viewed as an information resource themselves. Therefore, relations between an information resource and ontologies have to be created through explicit contextual information and specific references which are added to the information resource. This methodology clearly does not fit into any category on the knowledge management dimension other than Meta-context Approach.

The same reasons that lead to the obvious characterization of ontologies as a meta-context approach hamper the categorization with respect to technology forecasting. The question arises, which steps and activities of the technology forecasting process benefit from the development and application of an ontology. Still, one of the most essential requirements for a knowledge management technology in order to support technology forecasting and improve its efficiency, is the capability of facilitating the handling of large amounts of information and the identification of valuable pieces of information. In the sense of a semantic web one could argue that an ontology of a company’s knowledge domain could help to identify valuable information on the web through the application of semantic software agents. However, this is not yet a reality and therefore not applicable. If a company still wanted to apply a similar process, information resource annotation would be the most challenging activity. Automated methods to accomplish the annotation task, like the so-called method SemTag (Dill et al. 2003), are still in early stages of development. Without automated mechanisms to annotate pieces of information for the use within the company, it would be the task of individuals to
identify such pieces, analyze them, and finally annotate them. Without a doubt, such a process requires extended efforts and its quality depends on single individuals to a large extent. Nevertheless, the result of this process facilitates discovery and thereby reuse of the annotated information later on, especially in those cases where over time, the needed information can be found more and more inside a company’s information repositories.

Taking the nature of the information which is processed in technology forecasting for strategic innovation management into account, it is doubtful whether needed information will be found annotated inside company repositories. The efficiency of technology forecasting for strategic innovation management depends on the identification of valuable pieces of information, which lie inside the technological scope of the company or influence it, from a myriad of new information appearing every day in scientific articles, research reports, or other media. Thus, valuable information can be found more likely outside the company than inside. This is especially true for Step II – Obtain Information and Step VIII – Ongoing Monitoring for both of which the collection and identification of new information is essential.

Following the premises of Zelewski as presented in 4.2.5.2, possible applications for ontologies within technology forecasting can still be derived. Zelewski argued that the knowledge intensity of the tasks which are to be accomplished and the degree with which knowledge backgrounds of involved parties differ both influence the importance of ontologies as a means to improve efficiency of the considered process tasks. It has already been shown in 3.1.1 that technology forecasting itself can be seen as a knowledge creating activity. Moreover, when conducted for strategic innovation management, many technology forecasting methods, like for example the Delphi method, have the purpose to transfer certain knowledge of individuals into statements about future technological innovations and developments. Furthermore, an additional transfer of this knowledge takes place when the results of a forecast are applied to develop innovation strategies. In this case, the knowledge about the future has to be transferred from the individuals involved in the forecasting process to the managers responsible for the innovation strategy. Thus, the first of Zelewski’s two premises is true with regards to technology forecasting for strategic innovation management.

The second premise aims at the different knowledge backgrounds of the individuals involved. Figure 3-1 in chapter 3.1.1 illustrates the possible roles which are involved in the technology forecasting process. Most of the individuals which are incorporating
those roles come from inside the company. Moreover it can be assumed that such individuals come from a small number of departments and therefore have similar knowledge backgrounds. The reason for this is that departments like control, human resources or marketing are, assumingly, most often not involved in forecasting activities. Rather, departments like R&D, production, and the corresponding management levels are integrated into the technology forecasting process. However, the knowledge background of the latter departments will not differ to a large extent, due to the close interaction and the comparable field of topics of these departments. Only the communication and coordination with external individuals, which are involved in the forecasting process, involves differing knowledge backgrounds. Therefore, ontologies can be applied to facilitate communication and coordination with in the forecasting process involved parties from outside the company. Such parties are, for example, certain domain experts, universities, research institutes, or other companies.

As a result of the discussion above, only the category Step IV – Implement Methods seems to be suitable for the characterization of ontology application within technology forecasting for strategic innovation management, limited by the technology forecasting methods chosen.

4.3 Integrative Discussion of the Evaluation

At this point of the thesis, the knowledge management technologies data mining, information retrieval, case-based reasoning, topic maps, and ontologies have been analyzed with respect to their degree of context explication on the one hand and their potential in supporting technology forecasting on the other. Figure 4-13 illustrates the completed Context-Complexity Matrix which results from the discussions above.
Figure 4-12 Completed Context-Complexity Matrix

Obviously some steps of the process can be supported by more than one knowledge management technology. Therefore, the question arises, which single technology or which combination appears to be the most promising way to support and improve efficiency of the technology forecasting process for strategic innovation management?

To answer this question it is helpful to consider technology forecasting for strategic innovation management with respect to the type of input each process step requires. Section 4.1.1 has shown that the complexity of the information structure within the technology forecasting process increases in the course of the process. While at the beginning of the process, or more precisely during Step II – Obtain Information, the objective is to identify the needed type of information which corresponds with the scope and time horizon of a forecast, Step VIII – Ongoing Monitoring requires the discovery of specific information which closely related to the information structure created from forecasting results and the information used within the forecasting activities and allows evaluation and verification of forecasting results. One could argue that in the course of the technology forecasting process for strategic innovation management, context becomes more and more important. At the beginning of the process the importance of context as well as the information structure complexity is rather low, but in the end the degree of complexity and context importance reaches its maximum. As a result, the
Context-Complexity Matrix has to be refined to integrate the focus of strategic innovation management in such a way that only the upper left triangle represents possible solutions which are promising ways of supporting technology forecasting steps through knowledge management technologies. Figure 4-14 incorporates this idea.

![Figure 4-13 Revised Focus within the Context-Complexity Matrix](image)

A striking point of the Context-Complexity Matrix is the fact that topic maps are capable of supporting each process step in a certain way. However, topic maps require some knowledge about the domain and its topics for their generation, while information retrieval provides functionalities which require less prior knowledge and can be used to gather a first broad variety of information. This can be especially helpful during the first phases of technology forecasting research efforts. Such information can then be analyzed to generate the needed topic map which corresponds to the scope of a technology forecast. Later on, this can be used to classify and organize further information and hence allows a more systematic way of discovering additional information.

As a summary it can be stated that a knowledge management system which is based on topic map technologies and integrates information retrieval functionalities as extensions to those provided by the topic map, is the most promising solution for supporting technology forecasting for strategic innovation management.
5 A Real World Example

In order to illustrate the results of the evaluation and link them to a real world problem, an example process will be described and evaluated in the subsequent section after which a solution concept which applies the insights from the previous chapters will be developed. The example is taken from a project at DETECON, Inc. in San Mateo, USA, a technology and management consulting company with a focus on innovation engineering. The project is conducted for Deutsche Telekom AG with the main objectives being the identification of technology trends and developments with the ability to open new opportunities and assessment of their innovation potential. The goal of this is to give assistance with long term strategic innovation planning and improve Deutsche Telekom’s competitive advantage.

5.1 Example Scenario: DETECON, Inc. – Deutsche Telekom AG

The following two sections will focus on the technology forecasting process as it exists between DETECON, Inc. and Deutsche Telekom AG. The process will be described and compared to the technology forecasting process for strategic innovation management as presented in this thesis which will then be followed by the evaluation of the process at DETECON, Inc. in section 5.1.2.

5.1.1 Business Process Description

When compared to the general process of technology forecasting for strategic innovation management as presented in section 2.2.3, the process at DETECON shows several differences. First of all, people at DETECON distinguish between inductive and deductive approaches toward innovation and trend identification which is similar to the distinction of “core technologies” and “white spaces” by Reger as mentioned is section 2.2.3. Inductive methods begin with the identification and formulation of a certain problem setting. In this case the term problem refers to certain needs which emerge from inside Deutsche Telekom and are derived from internal developments, processes, and a like. After problem formulation, specific information is obtained which is related to the problem and helps find innovative solutions.

In other cases the appearance of information about certain technology developments and innovations precedes the identification of a problem. Emerging trends and innovations are monitored by DETECON and assessed with respect to their potential influence on Deutsche Telekom’s business or innovation strategy. An example for technological
trends which lead to an innovation need from outside Deutsche Telekom is Voice-over-IP which, in the long run, can be considered a threat to Telekom’s ISDN landline product. In summary a forecasting process at DETECON can be initiated in basically two ways: either problem precedes information or information precedes problem.

The process as developed within this thesis continues with steps III – V, Select Methods, Implement Methods, and Evaluate Methods. DETECON applies a multidimensional set of methods for information gathering for forecasting activities which ranges from formal and informal personal networks, literature research, as well as market and trend analysis to patent analysis, Delphi surveys, and cooperation with university, governmental and corporate R&D labs and standards organizations. A comparable set of methods exists for technology evaluation and assessment purposes. The selection of those methods which are applied in a certain case varies and depends on the technology type and the analysis depth needed.

Furthermore, the technology forecasting process for Deutsche Telekom can be considered an iterative process. A first iteration gives a broad overview of potential interesting technologies which are then communicated on a very high level towards Deutsche Telekom. On the base of interest Telekom requests an additional iteration with an increased analysis depth and a more precise technology scope which results in the development of so-called technology profile documents. The process leaves DETECON in a phase comparable to step VI, “Prepare Decisions”, of the general process of technology forecasting for innovation management. From this point on, Deutsche Telekom is responsible for the remaining steps and activities and the integration of the acquired knowledge into its innovation strategy.

In general, the motivation of the whole process is to provide ideas and analysis about future technology and business trends for Deutsche Telekom and thereby assist them in formulating innovation strategy and corresponding objectives. Figure 5-1 illustrates the technology forecasting process at DETECON, Inc. with respect to the Deutsche Telekom project.
Figure 5-1 Technology Forecasting Process at DETECON, Inc. for Deutsche Telekom AG

5.1.2 Process Evaluation

The technology forecasting process at DETECON is a good example of the fact that real world processes are likely to differ from generalized processes as found in scientific literature. The specific situation of conducting a service to deliver information and transfer knowledge from DETECON to Deutsche Telekom and thereby account for the requirements and needs of Deutsche Telekom necessitates more flexibility with respect to forecasting scope, time horizon, and even the forecasting process itself. It is obvious that DETECON’s process of technology forecasting comprises more dynamics than the process presented in section 2.2.3 of the thesis. Due to the bipartite initiation of the process which can also occur in parallel, the scope of the process not only varies, but different views and ranges of one scope or even several different scopes are needed or merged in order to achieve a comprehensive technology forecast. While an inductive approach is based on a scope defined with respect to the problem setting, the deductive approach applies a broad scope in order to detect all technological developments as early as possible and derive their consequences for Deutsche Telekom. In such a complex situation it is necessary that information is organized in such a way that it can be easily stored, related to other information, and brought into context in order to be found, recovered, and reused when needed.

This fact becomes even more important with respect to a second aspect which integrates additional dynamics into the process, i.e. the uncertain requirements and demands of Deutsche Telekom. More precise technology studies are not required for all of the identified technologies and are requested on a base of interest and relevance. However, the identification of a potential innovative technology and its assessment with respect to relevance for Deutsche Telekom already entails the collection and analysis of a certain amount of information. In case a more precise technology study is requested, it can be
assumed that a lot of the information needed has already been analyzed and certain knowledge already exists. The remaining information-need targets very specific information to extend the available information base. Therefore, to have an efficient way to recover information and its related items, such as analyses or contacts, and furthermore, to be able to find specific information selectively is a means to cope with the dynamics of the process and improve the overall efficiency of technology forecasting at DETECON, Inc.

Another consequence of the dynamics incorporated into the technology forecasting process at DETECON is the need to find efficient ways for the knowledge transfer from DETECON to Deutsche Telekom. The success of the project relies on the success of Deutsche Telekom’s innovation strategy to a certain extent. This is because the results of the activities on the side of DETECON are supposed to be direct influences on strategy formulation. Therefore, comprehensive knowledge transfer is the enabler for the perception of a high project quality and thereby a key aspect for the success of the project. In an optimal case, the same concept which supports knowledge management inside DETECON and facilitates coping with the dynamics of the process, also fosters knowledge transfer from DETECON to Deutsche Telekom.

Overall, the process at DETECON depends on the capabilities of individuals to analyze information to identify business and technology trends as well as future innovations. One can presume that this task becomes more and more difficult over time with respect to decreasing product life cycle lengths and shorter innovation intervals. It is a common assumption that today the size of the knowledge of mankind doubles every two to three years. Since innovation and technology forecasting operates at the forefront of knowledge development, one can expect a fast increase of the amount of considered information which leads to more time spent finding the right information and less time analyzing it. Therefore, external developments can be considered as another aspect which makes an efficient process support essential for successful technology forecasting activities with respect to the task conducted at DETECON.

### 5.2 Solution Concept

The description and evaluation of the technology forecasting process at DETECON revealed certain factors which lead to specific needs for knowledge management above the general needs for knowledge management within technology forecasting as explained in 3.1. The subsequent sections will describe a solution concept which
integrates the insights developed in the previous chapters in order to support technology forecasting at DETECON and provide ideas to improve and sustain efficiency. At first organizational issues will be described briefly after which knowledge management technologies will be selected and a system concept developed.

### 5.2.1 Organizational Issues

During the introduction of knowledge management in chapter 2 the importance of organizational methods for knowledge creation and dissemination has been explained. The SECI model of Nonaka and Takeuchi has been described in section 2.3.3.2 as one example of models explaining organizational knowledge creation. The author believes that such organizational aspects are relevant for the success of the technology forecasting projects at DETECON. The efficiency of the knowledge transfer from DETECON to Deutsche Telekom as stated previously can be influenced positively through the adaptation of concepts like the SECI model and thereby creating a structured proceeding of knowledge dissemination.

With respect to the existing process at DETECON the SECI model can be used to determine those positions in the process which are most useful for direct meetings and information and knowledge exchange between DETECON and Deutsche Telekom. After the complex initiation of the process through either inductive, deductive or even both methods, at least one cycle of socialization, externalization, combination, and internalization can be assumed to be completed when considering problem formulation as socialization among DETECON and Deutsche Telekom employees as well as externalization through the description of a certain problem setting. In the case of an inductive process, the next step Obtain Information can be characterized as combination and internalization through the collection of information and its combination with existing knowledge by DETECON employees. As a result, a good position for a personal meeting between DETECON and Deutsche Telekom individuals is before the subsequent step Select Methods. Following this procedure, additional positions for direct communication among DETECON and Deutsche Telekom employees besides the steps Present Results and Prepare Decisions can be identified.

The author believes that such organizational aspects are one important factor for successful knowledge transfer. However, a more comprehensive organizational concept requires a deeper organizational analysis which lies outside of the knowledge management technology focus of this thesis.
5.2.2 Knowledge Management System Concept

The evaluation of DETECON’s technology forecasting process pointed out that a knowledge management system for supporting the process has to be capable of coping with a certain amount of dynamics. In addition to that, the system is required to facilitate information organization, reuse, and furthermore, assist knowledge transfer. Because of the results of chapter 4, a system based on the central utilization of topic maps is most promising to be applied successfully to cover the existing needs and improve technology forecasting efficiency. The remainder of this section will describe such a system concept and the reason which make it most promising for improving process efficiency.

One characterizing aspect of technology forecasting at DETECON is the flexible scope which different steps and activities inside the process require. Topic maps are a knowledge management technology which can be tailored to suit such a flexible use. Section 4.2.4.3 already gave an idea about how topic maps can be used to represent the scope of a single technology forecast. In case of DETECON’s process a solution can be the development of one single comprehensive topic map. This topic map represents the basic structure of the domain knowledge applied as, for example, technologies and their relations and influences. Sophisticated methods can then be used to restrict this topic map to the necessary range for single activities. This is sufficient due to the reason that all forecasting activities of DETECON, Inc. deal with technology and innovation developments and their influences on Deutsche Telekom’s technology and business situation. Therefore, the basic domain remains the same and only the range of the considered section of the domain varies. A topic map which is build and maintained for the corresponding domain and which can be tailored to represent only subparts of the available information through the exploitation of the scope attribute of topic maps provides an efficient solution for the requirement of flexibility.

Obviously, the nature of a topic map also facilitates organization and reuse of information and therefore fulfills another requirement with respect to technology forecasting at DETECON. Information which has been used once can be stored in a repository and can be accessed through the topic map. It is also associated with analyses, contacts, or other related information. Therefore, knowledge structures once generated can be represented by the topic map and recovery of such structures is facilitated.
In addition to that, a topic map can be used to categorize new information by determining the topics which occur in the new information. This functionality can be combined with methods for automated retrieval of information. The information is retrieved from some source (most likely within the WWW), it is analyzed with respect to the occurring topics, and then added to the information repository. This process facilitates the identification of valuable new information without the need to analyze all new available information manually. Because of the fact that the information is available through the topic map, it can be accessed when needed. Some examples for such information are online magazines, research articles, and all kinds of reports.

Another requirement of a system for the support of the technology forecasting at DETECON is the facilitation of knowledge transfer. It has been mentioned that in the existing process, once a technology is considered interesting and relevant through Deutsche Telekom, a technology profile is created which is then sent to Deutsche Telekom. Additional personal or phone conversations about these technology profiles can be seen as the methods for knowledge transfer. However, one can assume that the efficiency of this proceeding is limited to a small number of technologies at a time, because of limited receptiveness of human beings. A solution for this problem is the integration of the mentioned profile documents into the structure of a topic map and their association with main topics and further relevant information about the corresponding technologies. In such a setting, less information has to be conveyed during personal meetings. Rather, the technology related knowledge can be transferred with the help of the topic map by allowing access to the profile documents and its related information. The meetings can then be used to discuss technology and business consequences and thereby create additional knowledge which goes beyond the technology itself.

The challenge of such a system, however, is the maintenance of the topic map. A fully manual maintenance implies the awareness of new developments. In other words, new technologies have to be known before they can be integrated into the topic map. Therefore, methods have to be found which facilitate this task by suggesting new topics and associations. Statistical methods as applied within automatic indexing or text mining can provide a useful starting point for the solution of this problem.

In summary, topic maps provide the needed degree of flexibility, facilitate information organization and reuse as well as knowledge transfer. Therefore, a system which bases
on topic maps can be considered as a solution to the increasing difficulties related to technology forecasting at DETECON. Figure 5-2 summarizes the components of the suggested topic map system concept.

![Knowledge Management System Concept for DETECON, Inc.](image)

**Figure 5-2 Knowledge Management System Concept for DETECON, Inc.**

### 5.3 Discussion

The example above shows that besides the need for knowledge management in technology forecasting processes as developed in 3.1 additional factors influence the usability of knowledge management technologies in real world situations. In the case of DETECON, because of the fact that technology forecasting is conducted as a service for another company, flexibility of the forecasting scope is essential. Priorities are much more dynamic, because DETECON has only indirect influence on Deutsche Telekom’s innovation strategy and therefore less control over the forecasting focus. It has been shown, that topic maps are a means to deal with such dynamics and that a system which is based on topic maps facilitates many process tasks and increases overall process efficiency. The resulting system concept shows how the insights of chapter 4 can be used to derive solutions for real world situations.
6 Conclusion

The final chapter of this thesis will give an outlook about where to go from this point on. Further leading questions are raised and ideas for additional research given. Subsequently, the whole thesis will be summarized and the main insights stressed.

6.1 Outlook

This thesis dealt with knowledge management technologies for the support of technology forecasting within strategic innovation management. Therefore, mainly the process of technology forecasting and its several steps were focused. From this point on, one could analyze whether knowledge management technologies are also capable of supporting single technology forecasting methods. Within innovation management, most forecasting is done via the analysis of information as shown by the example of DETECON and Deutsche Telekom. Integration of other forecasting methods into the supporting system as, for example, extrapolation methods, could lead to a higher forecasting quality and decreased uncertainty with the aim of automating a major part of the forecasting process and achieving improved decision support.

As mentioned in section 5.2.1 efficient knowledge management also depends on organizational issues to a certain extent. While this thesis considers knowledge management technologies to be central aspects for knowledge management support for technology forecasting, one could ask, what influence organizational knowledge management concepts have for technology forecasting. The author assumes that organizational concepts depend on the structure of the forecasting process. Processes which are conducted completely inside a single company might benefit more from organizational knowledge management concepts than processes which a scattered over one or two companies. Further research in this area could discuss which combination of technological and organizational process support results into the highest value for competitive advantage and company success.

Since this thesis is based on theory, another field of future research could be the transfer of thesis’ insights to real world applications. Such applications can then be used to measure the efficiency gain achieved through the integration of knowledge management technologies and thereby practical verification can be conducted.
6.2 Summary

The starting point of this thesis has been today’s situation of companies and their need to make decisions about future strategies and developments in shorter time frames and under increasing market pressure. Therefore, they apply technology forecasting as one means to reduce uncertainty and to develop innovation strategies with the aim to sustain or improve competitive advantage.

After the introduction of the three fields strategic innovation management, technology forecasting, and knowledge management, it has been shown that several factors lead to an increasing need of knowledge management support for technology forecasting within strategic innovation management. The efficiency of knowledge transfer from the forecasting team towards decision-makers has to be improved to increase decision quality. Moreover, technology forecasting can involve many individuals with varying roles within the process. Therefore there is a need for efficient information and knowledge dissemination among individuals involved in order to achieve comprehensive forecasting results. However, the most important factor suggesting the necessity of knowledge management support for technology forecasting within strategic innovation management is the increasing amount of available information about today’s advancements, technologies, and innovations. Not only does the amount of available information grow with increasing speed, but also the number of actual innovations and existing innovation fields grow at a comparable rate. Therefore, efficient identification of valuable information as early as possible is the driving force of knowledge management integration into technology forecasting.

The development of the Context-Complexity Matrix and its application on selected knowledge management technologies has shown that within technology forecasting an increasing amount of information structure complexity leads to an increasing need for context explication. Information repositories are less useful without the application of explicit meta-contexts which facilitate the discovery of needed information. While technologies like data mining or case-based reasoning provide only marginal efficiency increase, the concept of topic maps possess a broad applicability and has the potential to increase efficiency greatly, especially with respect to technology forecasting within strategic innovation management. This idea is furthermore developed through the DETECON, Inc. – Deutsche Telekom AG example which resulted in the creation of a
knowledge management system concept for support of the technology forecasting process at DETECON.

At this point, real world tests are needed in order to verify the theoretical results and ideas of this thesis and to identify further aspects with the potential of increasing technology forecasting efficiency, improving innovation strategy formulation, and thereby creating and sustaining competitive advantage.
Bibliography


Fayyad, U. M.; Piatetsky-Shapiro, G.; Smyth, P.; Uthurusamy, R. (1996): Advances in Knowledge Discovery and Data Mining, American Association for Artificial Intelligence, Menlo Park, California, USA


O'Hare, M. (1988): Innovate!: How to gain and sustain competitive advantage, Basil Blackwell Ltd., Oxford


