ABSTRACT

BAHETI, PRASHANT PRAKASHCHANDRA. Evaluating a Software Engineering Knowledge Base. (Under the direction of Dr Laurie Williams)

Often, a large amount of knowledge leaves the organization when an employee and the knowledge he or she possesses leave the organization. Knowledge Management (KM) facilitates the retention and distribution of intellectual capital within an organization and helps the organization to gain competitive advantage and create business value. Organizational learning occurs when knowledge from all projects is documented and stored in a repository to aid future decision-making. Such an information repository is called a Knowledge Base (KB). A KB is an important element of any KM system.

This thesis has the following objectives:

1. to describe the re-engineering of a software engineering KB within an organization;
2. to identify the critical elements that influence the success of a software engineering KB within an organization based on a case study at ABB; and
3. to propose a metric suite for evaluating a KB and to partially validate this metric suite using a case study.

This thesis proposes a set of metrics for evaluating a software engineering KB. The main goals of evaluating a KB are to determine whether the presence of the KB has benefited the organization (ABB) and to identify areas for improvement of the KB. The KB evaluation can be seen in the context of the Goal/Question/Metric (GQM) model. Via GQM, we set goals, refining the goals to a set of questions, and answer these questions through a set of metrics.

ABB had an existing Experience Database (ABB-1) containing experiences submitted by the employees as well as training materials. However, ABB-1 was not used by the

employees as intended because of certain deficiencies. As part of this research, these deficiencies were identified within the context of the KM reference model proposed by Abou-Zeid. This information was useful in the re-engineering of the ABB-1 to the ABB Software Process Initiative KB (ABB-2). The Quality Improvement Paradigm guided the re-engineering process.

The metrics collected indicate that ABB-2 is being presently used by only a small number of users. Although four out of five users felt that ABB-2 made it easy to find useful information, all the five users believe that ABB-2 produces useful information less than 60% of the times. However, the projected timesaving because of presence of ABB-2 are much greater than the time invested in developing and evaluating ABB-2, even by taking conservative estimates.

In our research, we have found four factors that are instrumental to the success of a KB.

1. assigning formal responsibility of the maintaining the KB to one or more persons;
2. proper publicity of the KB within the organization;
3. ability to easily find resources on KB, making it conducive to visit the KB for knowledge needs; and
4. employee participation in contributing to and regularly updating the KB.

The ABB case study can serve as a guide to implementing a KB within an organization. Partial validation has been performed for a subset of the proposed metrics. Since data collection mechanisms are still in place, a complete validation of the metrics will be completed after ABB-2 is regularly used by large number of employees.

EVALUATING A SOFTWARE ENGINEERING KNOWLEDGE BASE

by

PRASHANT PRAKASHCHANDRA BAHETI

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

COMPUTER SCIENCE

Raleigh

2003

APPROVED BY:

____

Committee Member	Committee Member

BIOGRAPHY

Prashant Prakashchandra Baheti completed his Bachelor of Engineering degree in Computer Engineering from Veermata Jijabai Technological Institute (VJTI) in Mumbai, India. After graduating from VJTI in June 2001, he joined North Carolina State University in Fall 2001 to pursue his Master of Science degree in Computer Science. During his Masters studies, he completed a summer internship in ABB, which entailed designing and developing a Knowledge Base containing product development resources for the development units of ABB.

ACKNOWLEDGEMENTS

I would first like to thank Dr Laurie Williams, whose continuous support and guidance helped me complete my thesis. I would also like to acknowledge my thesis committee members Dr Ana I. Antón and Dr Aldo Dagnino who gave useful input while I was working on my thesis project. Mr. Andrew Cordes from ABB worked with me on the Knowledge Base and provided valuable support and guidance. Mr. Eric Buccholtz was very helpful in deciding questions for personal interviews, and he guided me with his experience within ABB. Finally, I would like to thank North Carolina State University and ABB Inc., where I was given an opportunity to pursue my research interests. ABB also provided the funding and resources required to complete my research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Evaluating a Software Engineering Knowledge Base

1. Introduction

Organizations need to identify, manage, and exploit their knowledge assets [1]. Often, a large amount of knowledge leaves the organization when an employee and the knowledge he or she possesses leave the organization. Additionally, the knowledge within an organization should be available to whoever needs it so that repetitive work can be avoided. A software organization's main asset is its intellectual capital (IC). IC consists of the knowledge of the employees, the processes followed by the organization, and knowledge held by the organization regarding its customers [2]. Knowledge Management (KM) facilitates the retention, and distribution of this IC [3] and helps an organization gain competitive advantage and create business value [4]. A Knowledge Management System (KMS) is a set of tools and practices that support KM within an organization. One element of any KMS is an information repository of documented resources that can easily be disseminated to the intended audience. Organizational learning is enabled when knowledge from all projects is documented and stored in a repository to aid in future decision making. Such an information repository is called a Knowledge Base (KB), experience base, or knowledge repository, depending on its context and information it contains. Thus KM is a paradigm, a KMS is the actual system built by using the KM paradigm, a KB is a part of a KMS and the process of building the KB is a KM activity. Figure 1.1 shows the relation between KM, KB and KMS.

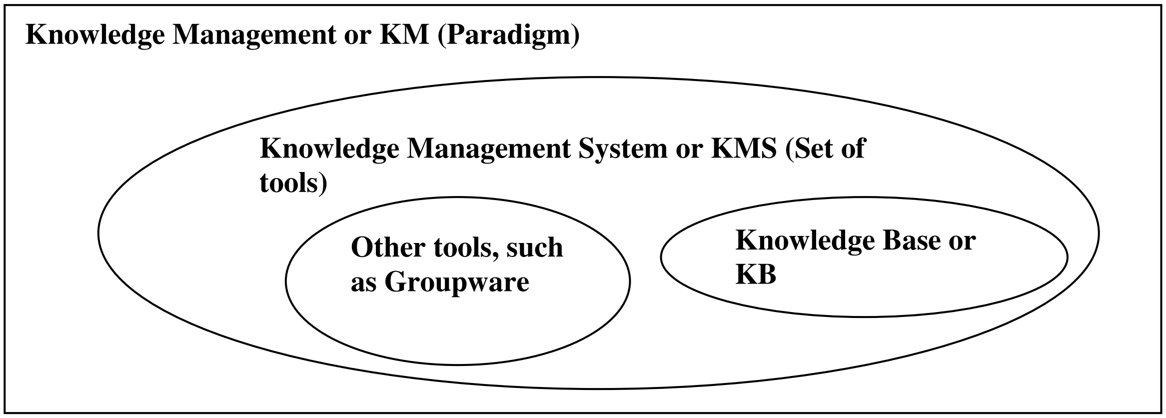

Figure 1.1: Relation between KM, KB and KMS

A KB is an important element of any KMS and can be used for document management within the organization [5]. For its success, a KB should be designed and populated such that the intended audience refers to the KB whenever a need arises. Once a KB is deployed within an organization, the organization should use suitable metrics to evaluate the efficacy of the KB. The use of predetermined metrics and targeted data collection allows organizations to determine if a KM activity is working as intended [6]. The analysis of these metrics can be used to guide the changes that need to be made in the design and deployment of the KB.

1.1. Types of Knowledge

Knowledge within an organization has two forms, tacit and explicit. Tacit knowledge is personal and context-specific and has not been documented or formalized [7]. In contrast, explicit knowledge is knowledge that has been documented in some form and can therefore be easily transmitted or distributed. Any organization would benefit from the ability to convert between these two types of knowledge to best exploit the knowledge within the organization. There are four types of knowledge conversions between tacit and explicit knowledge: tacit-tacit, tacit-explicit, explicit-tacit and explicit-explicit. KM involves all these types of knowledge conversions. Takeuchi and Nonaka [7] describe these four types of conversions as follows:

- *Tacit-tacit conversion (Socialization).* Socialization occurs when ideas are shared within a team and is a direct exchange of knowledge between people through verbal and visual interaction. These interactions could be discussions about projects and experiences, sharing views or brainstorming on problems [1]. Socialization helps one person to better understand how another person thinks.
- *Tacit-explicit conversion (Externalization).* Externalization is the process of codifying or formalizing the tacit knowledge held by people into some form that can be easily transmitted and distributed in a readable and understandable form. Some of the ways in which externalization takes place is through metaphors, analogies, hypotheses, or models.
- *Explicit-tacit conversion (Internalization).* Explicit knowledge is internalized when a person assimilates it. For example, a person may refer to a manual or a document to do a particular task, and thus internalize the knowledge obtained by reading the text. Documents and manuals help people in "re-experiencing" [7] the experience of others. Past experiences, such as others' success and failures stories, can help form a tacit mental model of how a practice is best performed within the organization.
- *Explicit-explicit conversion (Combination).* When existing knowledge is reconfigured or systemized through sorting, adding, combining, and/or categorizing, new knowledge can be created. Two similar documents may be combined to form a single document,

or a document may be restructured to make it more readable. Combination occurs when the one explicit form of knowledge is converted into another explicit form.

1.2. KM and SE

In the Software Engineering (SE) domain, KM can be defined as "a set of activities, techniques, and tools supporting the creation and transfer of SE knowledge throughout the organization" [8]. SE projects are often challenged by poorly defined requirements, frequent staff turnover, and volatile hardware and software platforms [9]. Software organizations possess knowledge in different areas; each knowledge area is critical for achieving business goals. Some of these areas include (1) documented knowledge about new technology, the organization's domain, and local policies and practices; and (2) the employees' tacit knowledge [3]. Organizations often face problems identifying the content, location, and use of their knowledge because it exists in different forms in the organization. Since KM can help alleviate some of these problems, many organizations realize that KM must be addressed in their process improvement initiative. Process improvement includes the provision of access to process-related artifacts that support SE and quality management techniques [8]. For example, software process knowledge can be externalized [7] by having a set of standard templates for different phases of development and support activities, such as a software requirements specification template.

1.3. Research Objectives

This thesis has the following objectives:

1. to describe the re-engineering of a SE KB within an organization;
2. to identify the critical elements that influence the success of a SE KB within an organization based on a case study of the implementation of a KB; and

3. to propose a metric suite for evaluating a KB and to partially validate this metric suite using a case study.

1.4. Case Study at ABB

The research objectives described in 1.4 were achieved by re-engineering a KB at ABB Inc.[1], based on the deficiencies of their existing KB. ABB is a power and automation technologies company that operates in nearly 100 countries and has approximately 135,000 employees. The author of this thesis proposes a set of metrics to evaluate a SE KB and uses the KB re-engineered by him at ABB to partially validate the metrics. Partial validation means that this thesis takes the first step in proving that the set of metrics proposed indeed measure what they purport to measure.

The rest of this thesis is organized as follows. Chapter 2 provides a summary of related work in KM. Chapter 3 describes the importance of KM metrics and previous work done with respect to KM metrics. Chapter 4 describes the metrics suite proposed in this thesis to evaluate a SE KB. The proposed metrics suite is partially validated using a case study at ABB. This case study is described in Chapter 5. The results and experiences from the case study are described in Chapter 6. Chapter 7 provides the conclusions for the thesis and future work required for validation of the metrics suite.

[1] www.abb.com

2. Related Work in KM

This chapter summarizes related work in KM. Section 2.1 describes KM initiatives in SE. Section 2.2 provides a summary of frameworks developed to oversee KM activities within an organization. Section 2.3 describes the Quality Improvement Paradigm (QIP) [10] that can be used in the evaluation of the processes used for a project. Section 2.4 gives a brief summary of the Goal/Question/Metric (GQM) [11-14] technique. Section 2.5 describes previous work in design rationale (DR) that may be used to guide the design of a KB.

2.1. KM efforts in SE

Several companies have initiated a SE KMS; we now discuss four such examples. One such initiative was at VTT Electronics [8]. A constraint of the VTT KMS was need for minimal impact on the software development and processes. Therefore, the solution could not require new technologies. Rather than solicit knowledge from software developers, VTT utilized software process improvement (SPI) experts as knowledge-capturing agents. The SPI experts gleaned information from project final reports, error databases, discussion forums, and through interviews with people in the organization. This information was packaged to provide easy access to customer projects information. This needs-based KM approach adopted by VTT worked well for one of the customer projects inside the company [8]. The customer project could retrieve the needed knowledge from the delivered knowledge package.

Infosys Technologies Ltd. is a software services company that manages organization-wide knowledge using three centrally-operated knowledge repositories [15]: the Knowledge Shop (K-Shop); the Process Asset Database; and the People Knowledge Map.

The K-Shop provides employees access to resources related to technology, domain, trends, culture, project experiences, and to internal and external literature through a web interface. The Process Asset Database captures "as-is" project deliverables, such as project plans, design documents, and test plans. The People Knowledge Map is an intranet-based system for employees to search and locate experts in different fields. KM has helped Infosys to increase its productivity by three percent because of effective reuse and to reduce defect levels by as much as 40 percent [15].

International Semiconductor Technology is an Integrated Circuit assembly and testing company where knowledge intensity, specificity, and volatility are comparable to those in SE and other knowledge-intensive fields. This organization utilized a KMS to assist with knowledge creation, update, sharing, and reuse [16]. Their KMS consists of a lessons-learned repository that contains validated knowledge, a case repository that contains disapproved cases or those undergoing validation, and an organizational directory. Knowledge creation is supported by soliciting experts for answering questions. The experts give their analyses, comments, and/or recommendations. The experts can discuss the problems using groupware until a solution is reached. The resulting solution becomes "validated" knowledge, which is stored in the lessons-learned repository. By supporting the sharing of critical, task-specific knowledge previously held by individuals, the company aims at developing better systems at a faster pace.

Basili et. al [17] are implementing a Commercial Off-The-Shelf (COTS) Lessons Learned Repository (CLLR) to support the Center for Empirically-Based Software Engineering (CeBASE) initiative. The CLLR contains journal articles, workshop presentations and government reports on COTS-based development. Each lesson has a set of attributes attached to it, such as the context in which the lesson was learned, type of data, recommended audience, relevant lifecycle phase. These lessons, available on the Internet[2], help managers to take decisions on COTS-based development projects. There is a mechanism for feedback on each lesson. As the CLLR evolves, new features such as ability to ask queries on particular topics, which will be answered by experts on the field, will be added. The dialogues between users and experts can be logged and inserted in the repository.

These four examples are representative of SE KMS that exist in different organizations. In each example, an important goal is to have easy access to both tacit and externalized knowledge. One or more knowledge repositories exist in each KMS. Another goal is to facilitate socialization [7] by creating an environment conducive to interaction between persons in the company. For example, whenever a person has a problem that can be solved by some other expert in the same organization, he or she should be able to find and contact that expert and obtain a solution.

2.2. KM Frameworks

Many organizations have built their own KM frameworks to oversee KM processes, methodologies, tools and techniques. A KM framework provides a set of principles, terms and concepts for guiding a KM effort within an organization. A KM framework takes into account factors such as:

1. strategic goals of the organization;
2. knowledge within the organization;
3. technological support within the organization;

---

[2] http://fc-md.umd.edu/ll/index.asp

4. learning through feedback loops; and
5. people/culture within the organization [18].

There exist two main kinds of frameworks: prescriptive frameworks and descriptive frameworks. Most of the frameworks are prescriptive in nature [18], implying these frameworks "provide direction on the types of KM procedures without providing specific details of how those procedures can/should be accomplished" [18]. Prescriptive frameworks are task-based because they specify the tasks such as acquiring, generating, organizing, sharing, and applying knowledge [18] without considering how KM fits into the goals of the organization, the people involved in KM activities, or the cultural context of the organization. Hence, prescriptive frameworks are unable to view KM holistically. Examples of organizations that have authored such prescriptive frameworks include Ernst and Young, Knowledge Associates, and The Knowledge Research Institute, Inc. [18].

Alternately, descriptive frameworks describe the procedures to achieve KM goals [18]. Descriptive frameworks may consider factors such as linking KM to business objectives, feedback loops within the organization, and cultural factors. However, even the descriptive frameworks have been unable to view KM holistically because there is no consensus as to which elements must be included in a descriptive framework [18]. For example, descriptive frameworks described by the National Technical Institute of Greece, Andersen Consulting, and The Delphi Group "include cultural factors in their framework but not learning or linkages with strategic business objectives." [18]

A holistic KM framework is both prescriptive and descriptive. Three example frameworks for KM are discussed below. Of these, Binney's KM spectrum is a prescriptive framework while the Knowledge Management Maturity Model and Abou-Zeid KM Reference model are holistic frameworks.

2.2.1. KM spectrum

Binney [19] developed a KM spectrum that helps organizations in understanding the range of KM options, applications, and technologies available to them, which in turn helps the organization to decide where to make its KM investments. Binney's spectrum is a KM framework that covers all the KM applications within an organization by categorizing them into six categories:

1. *Transactional KM.* In this case, knowledge is prepackaged and provided to the end user with the help of user interaction with the system. User interaction with the system can be termed as a transaction. Helpdesk and customer service applications are examples of transactional KM applications.
2. *Analytical KM.* Analytical KM involves the creation of knowledge from different sources of material. Data can be used to derive patterns or trends. Data warehousing and data mining fall into this category.
3. *Asset management KM.* Asset Management KM focuses on management of knowledge assets such as intellectual property or any other explicit knowledge asset. Document management and content management fall into this category.
4. *Process-based KM.* This is based on process improvement and codification of processes. Process-based KM is derived from other fields, such as process reengineering and total quality management.

5. *Developmental KM.* Developmental KM applications are used for increasing the competencies or capabilities of an organization's employees. Skills development and training applications fall into this category.
6. *Innovation/creation KM.* Applications belonging to this category provide an environment for people from different disciplines to collaborate in teams and create new knowledge. [19]

The KM applications described in the literature can be mapped to one or more elements of Binney's KM spectrum. Developing and deploying a KB within an organization, which is an objective of this thesis, is integral to document management. Thus a KB is a KM application which can be mapped to Asset Management KM. Binney's framework transcends most of the KM applications available and their enabling technologies, but it is prescriptive because it does not consider the relation of KM applications to the goals of the organization or how these KM applications affect the culture within the organization. Since this model did not fit the context of KM in ABB, it was not used to analyze the ABB case study.

2.2.2. Knowledge Management Maturity Model

The Knowledge Management Maturity Model (KMMM) [20], developed at Siemens AG, provides a holistic approach to KM. This model helps organization assess their current position with respect to KM and accordingly move ahead with new KM projects. The KMMM defines five maturity levels of KM, similar to those of the SEI's Capability Maturity Model (CMM) [21]. KMMM has three components:

1. *Analysis model.* The analysis model helps an organization analyze which key areas and topics should be developed in the future.

2. *Development model.* The development model describes how the key areas and topics can be best developed to reach the next maturity level. The maturity levels are the same as those defined for CMM:

   - *Level 1: Initial.* Organizations at Level 1 have knowledge generation, exchange, usage, and loss without any control.
   - *Level 2: Repeatable.* Organizations are conscious of the importance of KM activities for their business. Pilot KM projects exist because of motivation by certain individuals within the organization.
   - *Level 3: Defined.* KM activities are integrated into the day-to-day working of the organization and individual KM roles are defined and filled.
   - *Level 4: Managed.* The organization has "a common strategy and standardized approaches to the subject of knowledge management." [20] The efficiency of KM activities is regularly measured.
   - *Level 5: Optimizing.* The organization is capable of adapting to new requirements imposed by KM without dropping a maturity level. The measures provided by Level 4 coupled with strategic control help the organization to adapt to new requirements with respect to KM.

3. *Assessment process.* The assessment process is used to evaluate the maturity level of an organization with respect to KM. The phases involved in a typical KMMM assessment project are orientation and planning; motivation and data collection; consolidation and preparation; feedback and consensus; ideas for solutions and action proposals; and report and presentation.

KMMM provides a framework for an organization to understand the level of maturity of the organization with respect to KM and identify the specific key areas to develop to attain a higher level of KM maturity. Since ABB emphasizes business value more than reaching a KM maturity level, this model was not used for KM within ABB.

2.2.3. Abou-Zeid Reference Model

Abou-Zeid [22] composed a Knowledge Management Reference Model (KMRM) that is holistic and focuses on business value. The KMRM proposed by Abou-Zeid guides the development, operation, and assessment of a KMS. The model identifies the processes that need to be supported by any KMS. These processes change the state of existing knowledge within the organization to a state where knowledge exists in an updated, desirable, and accessible form. For example, one such process could be identifying the experts in an organization and documenting their views on specific problems to form new knowledge for future reference. Abou-Zeid's model identifies the processes that should be supported by a KMS through its three-layer reference model.

1. *Cognitive Domain Layer.* The model's first layer, called the cognitive domain layer, addresses entities, called business things (or *B-things*), that relate to business issues or organizational goals. An example of a B-thing is *efficient and cost-effective software production*. Each B-thing is associated with certain knowledge that enables or supports it. These are called knowledge things (or *K-things*). The "efficient and cost-effective software production" B-thing is supported through the reuse of the previously developed product development templates K-thing.
2. *Functional Layer.* In an organization, the B-things are relatively stable; however, the associated K-things are in a state of continual change as knowledge evolves. The

second layer, known as the functional layer, includes the processes required to change the state of the K-things; these are called the *K-manipulating* processes. For example, a K-manipulating process provides the ability for an employee or a development unit to contribute a template to a KMS. The processes supporting the K-manipulating processes are called *K-enabling* processes. The K-enabling processes produce cultural- and organizational-enabling conditions for the K-manipulating processes. Publicizing the existence of KM applications, and ways to use them within the organization can be termed as a K-enabling process.

3. *Resources Layer.* The third layer is the resources layer that consists of the enabling information/communication technologies (ICT) to support the K-manipulating and K-enabling processes. For example, the technologies that support different KM activities, such as a messaging application or organizational databases are a part of the resources layer.

The three layers, cognitive domain layer, functional layers, and resources layer and the elements of each layer are shown in Figure 2.1.

Cognitive Domain Layer (Layer 1) **B- things and K-things associated with B-things**
Functional Layer (Layer 2) **K-manipulating processes operating and K-enabling processes**
Resources Layer (Layer 3) **Organizational and Information/Communication Tools**

Figure 2.1: Elements in each layer of Abou-Zeid KMRM

The Abou-Zeid model identifies the tasks involved in KM as K-manipulating processes and the enabling factors for K-manipulating processes as K-enabling processes. Both are then mapped to corresponding B-things, and supported by the underlying ICT support. The Abou-Zeid KMRM provides a hybrid (descriptive and prescriptive) framework for KM systems, because it considers the cultural aspects as K-enabling factors and links the K-manipulating processes to organizational goals (B-things). Hence, the Abou-Zeid KMRM is said to be holistic. The Abou-Zeid model was published after the ABB KB had already been implemented, but it nicely fits the context of KM at ABB. Hence, this model was used for a retrospective analysis of the ABB Case Study and will act as guide for future work in KM in the organization.

2.3. Quality Improvement Paradigm

The Quality Improvement Paradigm (QIP) [10, 23] is a framework for corporate learning and improvement. QIP continually identifies organizational goals and assesses the status of the organization with respect to these goals. QIP defines six sequential steps for corporate learning and improvement [12]:

1. Characterizing the current project and its environment with respect to models and metrics based on knowledge about the organization's business. The environment for a project provides the context for defining goals, selecting a process to complete the project, evaluating the results, and predicting performance. This step sets the context while actual goal setting and process selection is done in subsequent steps.

2. Setting quantifiable goals depending upon the environment of the organization and the current project. These goals may be refined into a set of questions and metrics for successful project performance or improvement over previous projects.
3. Determining and tailoring a process model for the project, based on the context and environment of the organization, and the goals defined in step 2. This is the step where methods and techniques for the particular project are decided. The process must be well designed to allow data collection from the project for analysis in step 5.
4. Executing the processes and sub-processes. Additionally, data is collected for feedback about the progress of the project and take corrective actions, if required.
5. Analyzing the data to evaluate the project and current practices. The data can be used to determine problems and recommend improvements.
6. Packaging the experience from the current project for use in the next project.

At ABB, an existing experience database was re-engineered to a KB using the QIP.

2.4. Goal/Question/Metric Paradigm

The Goal/Question/Metric (GQM) [11-14] paradigm is used for setting project goals and defining them in a tractable manner. Each goal is refined into one or more quantifiable questions that will characterize the assessment or achievement of a goal [11, 12]. The data collected through a set of metrics is used to answer each question quantitatively. GQM is a measurement model with three levels:

1. *Conceptual level.* Goals are defined for objects of measurement such as:
    - products including artifacts, deliverables, and documents that are produced during the system life cycle; e.g. specifications, designs, programs and test suites.

- processes such as testing, interviewing, or designing that are normally associated with time; and
- resources such as personnel, hardware, software, and office space that are used by the processes to produce desired outputs.

2. *Operational level.* The questions for each goal are identified. These questions characterize the object of measurement (product, process or resource). The answers to these questions should be measurable through some set of metrics that can then be collected to answer the question, and, in turn, assess the goal.
3. *Quantitative level.* The quantitative level describes the metrics for each question identified in the operational level. The metrics must answer the question in a quantitative manner. The data for the metric can be subjective such as the level of user satisfaction or objective such as staff hours spent on a task. Even the subjective metrics can have quantitative results. For example, the level of user satisfaction can be a number ranging from 1 to 10.

This thesis uses the GQM model to identify the goals in evaluating a KB after it has been deployed within an organization. Each goal is then refined to specific questions and metrics can be used to answer these questions. The questions asked in evaluating a KB can be traced to the goals of the organization in investing on a KB.

2.5. Design Rationale

In software engineering one of the earliest forms of design knowledge captured to support software engineering was Design Rationale (DR) [24]. To understand the design of a system, it is useful to consider what different options are available for designing the system, which option is best, and why. DR captures the different options available for designing a system and the reasons for choosing a particular design [24]. DR provides an explanation as to why a designed artifact is the way it is by evaluating the options that were available for the design [25, 26]. Formally, the DR representation describes a design space that consists of a decision space and an evaluation space. A decision space documents the options available for a design, and an evaluation space documents the evaluation of these options, leading to the choice of a particular design. The evaluation space provides explicit reasons for choosing among the possible options. The effort of recording rationale should not be too burdensome for the designer and the DR repository should be sufficiently structured and indexed so that the documentation is retrievable and understandable to a person who needs to refer to the design to solve similar problems in the future.

Several approaches to DR have been described in literature; we outline a few of them. The origin of most approaches is the Issue-Based Information Systems (IBIS), which consists of Issues, which include anything that needs to be discussed, ways of resolving issues (known as Positions) and Arguments supporting or objecting each position. The Potts Bruns' Model consists of Artifacts representing objects of the design process. Artifact features raise Issues, which are resolved by Alternatives. Finally, each Alternative is supported or challenged by Justifications. The Questions, Options and Criteria (QOC) approach is similar to the Potts Bruns' model. The artifact features are called Options, and design issues raised by each artifact feature is framed as a Question. Criteria are used to choose between Options. Arguments may be attached to any artifact feature, in order to support or object an Option.[25]

Karsenty [27] conducted an empirical evaluation of DR documents with experienced professional designers. The study found that at least some designers extensively used DR as a support for understanding and assessing a previous design. These designers found DR to be useful but not sufficient. The study shows that less than half of the designers' DR related questions were answered by the DR document from previous projects. Hence, DR cannot be used as a standalone technique to guide design decisions; current information on projects is equally useful.

DR relates to this thesis in two ways.

1. DR is a method for creating an organizational memory [5] of why decisions have been made. Hence, it can be seen as KM activity. A KB can store the DR for various development projects, which can be put to future use. For example, Arango et. al. use technology books and product books to store design knowledge in the Schlumberger Laboratory for Computer Science [28]. Technology books contain the best knowledge available in the organization about a class of problems, such as data acquisition or measurement techniques. On the other hand, product books consolidate knowledge specific to individual products, such as version analyses and histories of deliberations (which is a part of most DR capture tools).[28]
2. When a KB is to be deployed within the organization, there are many options to design the KB. Hence, designing a KB itself can use DR principles. Before developing a KB, a design space can be used to identify the various options available to implement the KB. The options cover the structure of the KB, its interfaces, and the tools required to build and deploy it within the organization. Each of these options should then be evaluated in the context of the organization, and the best option can be chosen.

3. Metrics for Knowledge Work

Once a KB is deployed within an organization, the organization should choose suitable metrics to for its evaluation. Building a metrics suite for a KB requires an understanding of how knowledge can be measured. Knowledge is intangible in nature and, hence, difficult to measure. Additionally, KM is one of the many processes that impact the growth and culture of the company, making it difficult to assess the impact of KM on the organization in isolation. Hence, the metrics proposed to evaluate a KMS or particular tools that form a part of the KMS, are mostly indirect or approximate.

One goal of this thesis is to evaluate a KB. However, prior work in KM metrics has mostly considered the whole KMS, with the KB as an integrated part. The KM metrics related work described in this chapter served as a basis for devising the KB metrics suite. The rest of this chapter contains background work in measuring IC and impact of KM on an organization.

3.1. Importance of Metrics

The evaluation of a KMS helps determine whether an organization has been successful in achieving the goals it had set forth for the KMS.

> *Without measurement criteria, we are clueless as to whether we are accomplishing anything or not. Without measures, we turn KM into a metaphysical exercise with little actionable value to anyone.* [6]

An organization deploying a KMS generally has one or more of the following objectives [29]:

1. Avoiding costly mistakes;
2. Sharing best practices;

3. Faster problem solving;
4. Faster development times;
5. Better customer solutions;
6. Gaining new business; and/or
7. Improved customer service.

These objectives help organizations gain competitive advantage within the market and make their internal processes more efficient. To achieve these objectives, organizations strive to exploit knowledge within the organization in one or more of the following areas [29]:

1. *Customer Knowledge.* Customers provide valuable insights to the organization into the use of a product and the requirements that need to be met.
2. *Knowledge about products and services.* Organizations try to leverage the knowledge they possess about products and services provided by the organization.
3. *Knowledge in processes followed by the organization.* Knowledge in processes includes all process-related artifacts used and developed within the organization.
4. *Organizational Memory.* Organizational memory consists of company databases, procedures, and guidelines. It also includes employees' tacit knowledge.
5. *Knowledge in relationships.* Knowledge is held by the employees by virtue of working with each other, in pairs, in groups or with a particular customer.

A KB can be used to leverage the first four areas listed above.

3.2. Prior work in KM metrics

Metrics in KM can be measured at different levels of granularity. Granularity refers to the fineness with which metrics are collected. For example, there could be metrics that show how the presence of a KMS as a whole affects the financial situation of the company and other metrics that focus on the components of the KMS. These KM metrics are collected at the business level. Alternatively, it is possible to evaluate the effects of particular KM activities and tools within the organization at a finer level of granularity; these finer measures enable the organization to determine how effectively KM has been implemented and how each KM activity individually benefits the organization. The focus of this thesis is to implement a KB within an organization for document management, which is an important KM activity. Since the related work in KM metrics has not been isolated for a KB, established metrics for a KMS act as a guide to deciding the metrics for a KB.

Because of the evolving nature of knowledge, "the effect of using it better can never be measured as a simple one-to-one correspondence (so much knowledge in, so much product out)" [30]. Measuring knowledge involves measuring the IC added to the organization because of the KM effort. IC itself is intangible in nature and hence difficult to measure. Hence, indirect measures or indicators of IC are used for measurement purposes.

Market value of a company is the company's stock price multiplied by the number of outstanding shares, while book value is the value of the company's physical assets [2]. The difference between market and book value is the value the market places on the intangible assets of the company. One measure for IC is the Market-to-Book value [2] of the company. Market and book value can be measured before and after the launch of a KM initiative and be used to assess the value of the KMS. However, the measure is so abstract that it is not possible to isolate which actions affect the aggregate intangible assets [2] and how market factors come into play.

The impact of a KMS within an organization can be measured in terms of factors, such as project completion times, customer satisfaction, and employee productivity. These factors are indicators of the economic value produced by the firm's IC. The term *IC stock* refers to the intangible assets of the organization, such as the tacit knowledge of the employees and customer knowledge. The Effective KM Working Group is the outcome of a cross-industry effort in IC management. This group consists of members from seven large U.S. corporations namely Charles Schwab, Chevron, Dow Chemical, EDS, Motorola, Polaroid, and PriceWaterhouseCoopers. This group has produced a measurement framework that identifies a small set of indicators of IC stocks and measures of the economic value associated with them [2]. An organization can identify a set of appropriate core indicators of IC that are suitable in its context. A subset of these indicators can be used to evaluate the impact of a particular KM activity. Examples indicators include:

1. human capital indicators, such as retention of key personnel, employee satisfaction, employee commitment, and training expenditures measured as a percentage of payrolls;
2. innovation capital indicators, such as expenditures on research and development and percentage of workforce involved in innovative work;
3. process capital indicators such as percentage of business-critical processes documented and analyzed and the percentage of these documented processes are actually used; and
4. customer capital indicators such as customer satisfaction, customer retention, product or service quality and percentage of repeat orders [2].

These indicators can be measured objectively and/or subjectively. For example, the percentage of the documented processes actually used within an organization is an objective metric whereas customer satisfaction requires a subjective assessment as well. A KMS should facilitate the distribution and retention of an organization's IC. Hence, a relevant subset of these indicators can be measured before and after a KM project is deployed within the organization to assess the impact of the project within the organization.

Another method for measuring KM within the organization is benchmarking. Benchmarking can be defined as the "search for industry-wide best practices that lead to superior performance" [31]. The organization's KM initiative can be benchmarked against other units within the same organization, competing firms, and cross-industry organizations that are not competing with the organization [4]. An organization can measure its performance in some field or aspect with respect to the best in that area and try to emulate the organization that has proved to be successful. KM benchmarking is performed with respect to some performance areas such as speed of product development or the presence of a document management and groupware tool for knowledge sharing within the organization. Benchmarking can also be used to compare an organization's KM initiative with respect to other organizations, and if the ideal value of certain variables such as speed of product development is known, the organization can use this value to compare the effectiveness of its own initiative.

Lopez et. al. [32] deal with the impact of KM activities, as KM matures within the organization, and suggest metrics at each stage of maturity with respect to KM. According to Lopez, KM within a company undergoes five stages as discussed below:

1. *Enter and Advocate.* KM is formally conceived in the organization when it recognizes the need to manage knowledge. Someone or some group within the organization

demonstrates the value of KM and creates a vision for a KM initiative. Measurements should be used to expose the need for KM within the organization. These measurements can be used to identify inefficiencies within the organization, such as lack of reuse of resources, redundant development of artifacts, and loss of knowledge because it is either not documented or not accessible. KM needs are uncovered through interviews with the stakeholders (any one who holds a stake in the organization such as employee or customer).

2. Explore and Experiment. At this stage, knowledge-enabling activities are initiated in certain sections of the company to attract the attention of senior management. A knowledge-enabling activity is any activity that facilitates KM within an organization, such as educating co-workers about the potential benefits of KM and gaining their support. This is also the stage when teams that operate in an environment that fosters teamwork, information exchange and collaboration are identified.

Metrics at this stage are meant to explore ways in which KM can be implemented within the organization. There are anecdotal indicators such as success stories and lessons learned, along with qualitative and quantitative metrics. During the explore and experiment stage, benchmarking can be used to determine what the competitors and/or other divisions within the organization are doing to share knowledge internally and with its customers.

3. Discover and Conduct Pilots. At this stage, the KM initiative has been formally implemented by the organization. The goal of Stage 3 is to provide evidence of KM's business value by conducting pilots and capturing lessons learned that help the organization better implement KM on a larger and expanding scale [32]. Business value can be determined by comparing the time saved through pilot KM experiments versus the work involved, and identifying places within the organization where these pilot results can be duplicated.

An example of a KM pilot is to design and implement a KB within the organization. The following metrics can be used to measure the impact of a KB deployed within an organization:

- the amount of time spent, on average, by a user each time he visits the KB;
- number of repeat users; and
- the number of times a site is visited.

The cultural impact can also be measured through anecdotal stories about how employees feel about working in a team and collaborating with others. The metrics suite proposed by this thesis adopts metrics from Level 3.

4. Expand and Support. At this stage, KM is a part of the organization's funded activities and its deployment begins to become more widespread within the organization. Measures at this stage involves seeking answers to the following questions:

- Have the processes within the company improved as a result of deploying KM?
- Is KM properly executed within the company?
- Do managers support knowledge sharing?
- Does KM facilitate innovation within the organization?
- Is codifying knowledge benefiting the organization or is it too expensive?

5. Institutionalize KM. At this stage, KM is an integral part of the organization's "way of doing business." Full enterprise-wide deployment of KM is achieved. The KM measures involve checking progress of KM within the organization and monitoring how the KM culture evolves within the company. Lopez does not propose any specific metrics for this stage.

According to Vestal [33], KM measures can be classified as leading measures and lagging measures. *Leading measures* are those that provide feedback about the ongoing KM implementation and help in monitoring activities during the implementation as well as in adapting to new requirements. Example metrics to evaluate leading measure include the number of best practices shared, number of people with knowledge sharing as an objective, and the number of users accessing the KB. *Lagging measures* are those that show how the implementation has affected the business and are collected after the KM implementation is complete for a retrospective analysis of how KM has impacted the organization. The lagging measures reflect trends such as productivity improvements, time saved and revenue gain. This thesis adopts both leading and lagging measures.

3.3. Validation of Metrics

Once a set of metrics for any type of measurement is proposed, it is necessary to systematically validate them [34]. Validating a metric means providing convincing proof that:

1. a metric measures what it purports to measure, i.e. the metric is well defined and consistent with the properties of the attribute that the metric claims to measure;
2. the metric is associated with some important external attribute of the process or product, such as cost or maintainability; and
3. the metric is an improvement over existing metrics [34, 35].

There are two types of relevant validation for purposes of this thesis, theoretical or internal validation and empirical or external validation [34, 35]. *Theoretical validation* maps to point (1) in the list above, and involves clarifying the properties of the attribute to be measured, and analytically proving that the metric satisfies those properties. Such attributes are termed as internal attributes. Theoretical validation requires consensus among the research community regarding the properties of attributes [34] (e.g. effectiveness of a KB), and reaching such a consensus could potentially take many years.

Empirical validation entails demonstrating points (2) and (3) above. Empirical validation requires correlating the metric to the external attribute by comparing the values of the metric with the values of the external attribute. For example, a metric that measures the number of downloads for a KB may be related to external attributes of the KB such as awareness within the organization. Empirical validation would then entail measuring the number of downloads for the KB and determining whether the number of downloads is a realistic awareness indicator. Empirical validation also requires proving that a proposed metric is better than existing metrics in terms of factors such as ease of collection or relation to external attributes. It is important to note that empirical validation is very time consuming because it requires performing many studies to produce convincing evidence that a metric is indeed a measure of a given external attribute.

In a strict sense, it may not be possible to externally validate a metric, because it is tough to decide how much evidence is “enough” evidence. If many studies fail to invalidate a metric, then that can be treated as informal validation evidence of the metric. As the weight of the evidence increases, so does confidence in the metric. Increasing confidence in a metric is thus an indication that a consensus is being reached; however, such a metric cannot be validated per se. [34]

The following steps are required to validate the metrics for a KB:

1. collect data for the proposed metrics;
2. determine the actual effect of the KB on the external attribute; for example, monitoring the effect of the KB on live projects within the organization, and evaluating profit from using the KB;
3. determine how a given metric relates to the external attribute and whether a good value of the metric correlates to a good value for the external attribute; and
4. if possible, comparing the metric with other validated metrics in terms of effectiveness and ease of data collection. At this stage, the metric is said to be externally validated.

This thesis provides partial validation for a proposed suite of metrics for a KB deployed within a software development organization. It provides the first step for external validation of the previously discussed metrics. The internal attributes for a KB are usage, usefulness, effectiveness and ROI as explained in Chapter 4. The external attributes are profit from investing in a KB and awareness of the KB within the organization.

4. Proposed Metrics

The GQM (Goal/Question/Metric) paradigm is a structured approach for setting project goals, identifying questions (which if answered will satisfy the goals) and using metric(s) to answer each question. The process of evaluating a KB within an organization can be guided by the GQM approach. The main goals of evaluating a KB are

- to determine whether the presence of the KB has benefited the organization (profit); and
- to ensure that the potential users are aware of the presence of the KB (awareness) and find areas for improvement.

Thus, profit and awareness are external attributes of the KB for external validation. These goals can be refined into sub-goals i.e. the fulfillment of the main goals is contingent upon the following sub-goals:

1. to determine whether the KB is used by a large number of individuals or groups within the organization, which resources are most accessed and who accesses the resources(usage);
2. to determine whether the KB provides useful information to users (usefulness);
3. to determine how well the KB is designed and maintained to serve the needs of its users (effectiveness); and
4. to determine whether the benefits from the KB outweigh the effort involved in designing, implementing, deploying and maintaining the KB (Return on Investment).

These are treated as the goals in the GQM analysis of a KB evaluation project. The internal attributes of a KB are usage, usefulness, effectiveness and ROI. We refer to each of the sub-goals as the goals for GQM.

In the following sub-sections, our proposed metrics for measuring each internal attribute of a KB is explained. In each sub-section, the questions related to each goal are described, followed by the metric(s) that would answer that question. As we discuss in Chapter 6, a subset of these metrics is partially validated.

4.1. Usage Metrics

Usage is a measure of volume of the traffic on the KB and the types of resources most accessed. The usage metrics proposed by this thesis are given in Table 4.1:

Table 4.1: Usage Metrics

Goal	To determine whether the KB is used by a large number of individuals or groups within the organization, which resources are most accessed and who accesses the resources**.**
Question #1	What is the volume of traffic on the KB?
Metrics	• The number of times any page on the KB is accessed per week (also known as the number of hits) • The number of unique persons accessing the KB [32] • The total number of queries from the entire user population per week
Question #2	What types of resources should be available in the KB?
Metrics	• Contents of the KB that are most accessed • Subjective user feedback • User profiles (i.e. percentage of managers compared to percentage of developers)

Question #3	What is the geographic distribution of the users?
Metrics	Countries/cities are most actively using the KB?

Collecting usage metrics requires tools that can collect web statistics. The profiles of users accessing the KB and percentage of repeat users can be obtained by logging IP addresses. If any metrics are collected can identify a user, e.g. by logging IP addresses, the given organization's privacy policy should inform the user about how their personal information will be collected and used. For example, the users should be informed that their IP address will be logged and that their activity on the KB will be monitored to better evaluate the KB. The users should be given the option to "opt-out" of any data analysis that involves their personally identifiable information. Similarly, in personal interviews, the interviewees should be informed as to how their personal information will be used in data analysis.

Usage metrics quantify the traffic on the KB, indicating the number of individuals within the company benefiting from the KB. Usage metrics also provide valuable feedback to the KB implementation/deployment team for future modifications and enhancements to the KB. For example, knowing that certain contents of the KB are accessed the most could help in restructuring the KB to ensure that the frequently accessed resources are most easily available (e.g. by reducing the number of clicks to get to the resource). Knowing the profiles of users (what percentage of users are developers, managers?) can help in deciding the right content to be posted on the KB. A survey question can be used to determine what kinds of information the users would like to see on the KB. Geographic distribution of the users indicates which countries/cities most use the KB. Geographic data may also suggest the regions in which greater publicity and awareness of the KB is needed.

4.2. Usefulness Metrics

The KB's usefulness is measured to determine if the KB has fulfilled its basic purpose of facilitating access to resources. The usefulness metrics proposed by this thesis are given in Table 4.2:

Table 4.2: Usefulness Metrics

Goal	**To determine whether the KB provides useful information to users.**
Question	Does the KB aid the users in finding useful information?
Metrics	• Time per visit on the KB, where a visit is termed as the time spent interacting with the KB [32] • Average time spent per user on each page of the KB • Percentage of repeat users [32] • Percentage of artifacts downloaded and tailored • Subjective feedback from the users

User feedback can be obtained by means of a poll or a survey. A poll question can be presented to the user every time he accesses some resource to determine whether the resource obtained was useful. The percentage of artifacts downloaded and tailored by a development unit in a project can indicate how useful the KB is in live projects. This information can be obtained from the development units. The rest of the usefulness metrics can be measured with the help of a web analysis tool. Time per visit allows us to find out how much time is being spent on each page. If a page only contains links to resources, one would expect that much time be not spent on the page as the user clicks on a link to leave the page. This time should be typically less than a minute. If the average time for which such a page is viewed is unusually large (say more than two minutes), it would mean that users are not able to easily find the required links. However, if a page contains descriptive material, the user would take more time to view the page. If the time viewed for such a page is less, then it would could mean that either the users arrived at the page by mistake or the material posted on the page was not useful for the users. Either ways, it raises questions on how well the KB is structured and populated. Time spent on the entire site is an indication of how much time users spend on the KB, and an unusually low value of this could be a matter of concern. Percentage of repeat users indicates how many users come back to the KB, proving that they found something useful the last time.

4.3. Effectiveness Metrics

Effectiveness metrics measure how well the KB meets the users' needs, and whether the KB is kept up to date. The effectiveness metrics proposed by this thesis are given in Table 4.3:

Table 4.3: Effectiveness Metrics

Goal	**To determine how well the KB is designed and maintained to serve the needs of its users.**
Question #1	How often does the KB find relevant information?
Metrics	• Percentage of the resources on the KB downloaded at least once

	• User feedback to the question *How often does the KB find relevant information?*
Question #2	Does the KB facilitate contacting experts?
Metric	• User feedback to the question *Does the KB facilitate contacting experts?*
Question #3	How often is the KB updated?
Metrics	• Frequency of contribution per user; and • Number of new resources published on the KB per month

Effectiveness can be partly measured by user feedback through survey questions such as:

1. What percentage of times does the KB produce useful information? This metric has the limitation of being a very subjective measure and can only be used as an indicator.
2. What percentage of times were you able to obtain resources not present on the KB by contacting an expert?

All other metrics mentioned in this section require tool support.

Finding the percentage of resources downloaded at least once helps in finding if the right kinds of materials are being posted on the KB. Frequency of contributions and number of resources published on the KB indicate how often the KB is updated and what percentage of the contributions are actually published.

4.4. Return on Investment (ROI)

Once a KB is implemented and is in use within the organization, the organization should determine whether the benefits from the KB outweigh the effort involved. ROI (in the context of our research) is a measure of the economic value of investing in a KB within the company [2]. The investments made in developing and deploying a KB can be measured by summing the number of person-hours spent in developing the KB and accounting for technological investments, if any. However, it is difficult to measure ROI for a KB because of the intangible nature of IC. To determine the ROI in deploying a KB, we need to isolate the KB from any other factor affecting the financial state of the company. But, it is not practically possible to keep all other factors constant in the company. Hence, ROI measures are approximate.

This thesis proposes that ROI be measured through feedback from the development units, once the KB has been deployed and is in regular use. In assessing the ROI for a KB, a prime factor is the amount of work saved within the organization because of the presence of the KB. The ROI metrics proposed by this thesis are given in Table 4.4:

Table 4.4: ROI Metrics

Goal	**To determine whether the benefits from the KB outweigh the effort involved in designing, implementing, deploying and maintaining the KB.**
Question	Was the investment in the KB a profitable one?
Metrics	• Net person-hours saved per week because of reuse • Net person-hours saved per week because of easy access to resources and ability to contact experts

We now explain how the suggested metrics can be collected:

1. *Net person-hours saved per week because of reuse.* The different units within an organization may have to do a significant amount of repetitive work in the absence of an information repository providing access to key resources. This repetitive work could lead to many versions of the same type of documents present inside the organization. These documents include templates, guidelines, training materials, or any other resources that are not available through the information repository. The time saved because of reuse and the time that will be saved for further development within the organization can be quantified by determining the approximate difference in development and customization time of each type of document that exists on the KB, and multiplying this value by the number of documents that have previously been accessed from the KB or will be accessed from the KB in the near future (based on feedback from development units). For example, an approximate quantitative measurement of time saved in developing templates within the organization can be calculated as

   Time saved per template (NetTime) = Development time (developTime) – Customization time(customizeTime)

   Interviewing the development units could provide feedback about how many templates have already been downloaded from the KB and customized, and how many templates are yet to be developed. Knowing NetTime and the number of templates *N* already developed using the KB, an approximate value of the development effort in terms of person-hours saved because of the presence of the KB can be evaluated as.

   Total time saved per development unit (TimeSaved) = NetTime * N

Total time saved in the organization can be found by multiplying *TimeSaved* by the number of development units that will potentially access the KB. Knowing the number of artifacts that will be developed by the development units can help in projecting the number of person-hours that will be saved in the future. This figure can be obtained through feedback from the development units.

2. *Net person-hours saved per week because of easy access to resources and ability to contact experts.* The presence of the KB can offer users a single source for resources and, in the case of the ABB KB helps users contact experts in the field if required. Accessing a resource in the absence of the KB could take more time, because the user would have to search the resource from different sources within the organization. After the KB is being actively used by a sizeable population (which varies by organization), user feedback can indicate the number of person-hours saved within the organization because of the presence of the KB. Since this is a very subjective measure, this metric should be treated as an indicator of time saved, rather than treating it as an accurate metric.

4.5. Summary of metrics

The metrics and their respective proposed method of measurement are described in this section and are summarized in the Table 4.5. Usage, usefulness and effectiveness will be validated with both awareness and profit as the external attributes whereas ROI is directly related to profit.

Table 4.5: Summary of Metrics to Evaluate a KB

Metric Category	Metric	Proposed Method of Measurement
Usage	• Number of hits • The number of unique persons accessing the KB • The total number of queries from the users per week; • Contents of the KB most accessed • Profiles of the users • The frequency of accessing the KB per month • User feedback on what types of resources are required on the KB	Tool support, Survey
Usefulness	• Time per visit on the KB • Percentage of repeat users • User feedback on usefulness • Percentage of artifacts downloaded and tailored	Survey/Poll, Tool support

Metric Category	**Metric**	**Proposed Method of Measurement**
Effectiveness	• Percentage of the resources on the KB downloaded at least once • Frequency of contribution per user • Number of new resources published on the KB per month. • User feedback on percentage of times the KB produces useful information	Survey/Poll, Tool support
ROI	• Net person-hours saved per week due to reuse • Net person-hours saved per week because of easy access to resources and experts	Interviewing development units, survey

4.6. Analysis methods

Once the data for each metric is collected, the data needs to be analyzed relative to the expected results. As previously mentioned, the collected data may be either qualitative or quantitative. A web statistic such as the number of hits on the KB provides information as to how many people in the organization are actually using the KB, is an example of quantitative metric. However, user feedback on an open-ended question is qualitative in nature.

Each metric is analyzed depending on the nature of the data related to the metric. Quantitative data are analyzed through suitable statistical tests. For example, some metrics are obtained through closed-ended questions that have predetermined answers. The results of closed-ended questions obtained through the surveys/polls can be tested for statistical significance, using the Chi-square test. The Chi-square test is used to find whether the observed distribution is different from the expected distribution. Consider for example, a survey question such as:

What percentage of times does the KB produce useful information?

1. *< 20*
2. *20-39*
3. *40-59*
4. *60-79*
5. *> 80*

The above question has five possible answers and the Chi-square test can help us determine whether the observed distribution of responses for these five options is statistically significantly in comparison to the expected distribution. For example, consider that it is expected that 40% of the population will opt for option 5 and 15% will opt for each of the other four options. The actual distribution of responses is then obtained to test against this hypothesis. The expected distribution of responses and the actual distribution of responses are then used to determine the Chi-square statistic. This statistic can then be used to obtain a p-value to indicate whether the obtained and expected results are statistically significantly different.

In addition to closed-ended questions, open-ended questions may also be asked in a survey or interview. Qualitative data is obtained from such open-ended questions. Qualitative data can be analyzed by identifying themes, patterns or concepts [36] in the data. Meaningful information can be obtained from the data by categorizing the data depending on behaviors, phrases or interactions. These categories may be preset where the themes are fixed up front and the responses (qualitative data) are analyzed to match one or more themes. In contrast, the categories may be emergent i.e. the categories emerge as a result of the analysis. [36] Once the categories are identified, patterns are found within and between categories. The final results of the analysis describe these patterns and categories. For example, "*What other kind of information would you like to access through the KB?"* is an open-ended question whose results help in evaluating the KB as well as providing feedback for improving the KB to better tailor it to the needs of its users.

In the next chapter, the ABB KB case study is described. The chapter explains the re-engineering effort undertaken by the author of this thesis, the analysis of the re-engineering effort, and the initial data collected to partially validate the metrics suite proposed in this chapter.

5. ABB Case Study

ABB is a power and automation technologies company that operates in about 100 countries and has approximately 135,000 employees. The company has around 100 product development units, called business area units (BAU) that are responsible for software and product development. ABB also has a software process improvement group called the ASPI (ABB Software Process Initiative) group. ABB has pioneered its own product development model called the ABB Gate Model [37, 38].

The ASPI group discovered that, in such a large organization, there was potential for reuse of artifacts that are commonly used within the company, such as process-related templates, guidelines, and training materials. However, the BAUs of ABB had no access to a common source for these artifacts. Consequently, duplicate versions of these artifacts were either developed within the BAUs or the established versions took a long time to obtain. This led the ASPI group to deploy KM within their organization. The ABB KM implementation and re-implementation serve as the focus of the case study discussed herein; the research objectives are outlined in Section 1.4.

The first activity for the KM initiative at ABB was the implementation of a KB as a centralized resource repository for resources for development groups. The initial KB was called the Experience Database (henceforth referred to as ABB-1). The artifacts in ABB-1 were later moved to the ASPI KB (henceforth referred to as ABB-2) because ABB-1 was not being used as intended and had certain deficiencies. These deficiencies will be outlined in section 5.1.

The re-engineering of ABB-1 to ABB-2 can be analyzed within the context of the Abou-Zeid reference model. The six steps of QIP were used to guide to the re-engineering process, which we describe in this and the next chapter. Section 5.1 performs a retrospective analysis of ABB-1 in the context of the Abou-Zeid reference model. Section 5.2 sets the goals for ABB-2 based on the deficiencies identified in section 5.1. Section 5.3 outlines the processes for achieving the goals identified in section 5.2. Section 5.4 describes the execution of the processes and Section 5.5 describes the data collection mechanisms employed for ABB-2. Finally, Section 5.6 outlines the future work for ABB-2.

5.1. ABB-1: A Retrospective Analysis

In the first QIP step, the existing system is characterized i.e. the system characteristics and environmental variables are identified. ABB-1 contained experiences and training packages. However, the system lacked critical elements necessary for effective update and use by development units. ABB-1 consisted of a Lotus Notes database with a custom-built HTML interface. The database was a collection of documented software development experiences and good practices from ABB's BAUs, collected through the submittal of structured experience "packages." These packages underwent a multi-phased review and approval process before being accepted as entries to the database. The database provided a limited search capability of the experience packages and did not offer a mechanism for identifying pertinent external resources and web links.

Based on the feedback from the users and analysis by the ASPI group, it was discovered that ABB-1 had certain deficiencies. These deficiencies can be analyzed in the context of the Abou-Zeid KMRM [22]. The Abou-Zeid model describes business things (or *B-things*) as entities that relate to business issues or organizational goals. The B-thing for ABB is efficient and cost effective software/product development. This B-thing is associated with certain knowledge that enables or supports it, called knowledge things (or K-things). One K-thing related to the B-thing is the KB (ABB-1) with its artifacts. *K-manipulating* processes change the state of the K-things. The mechanism to contribute resources to the KB is an example of a K-manipulating process. *K-enabling* processes support the K-manipulating processes by producing cultural and organizational enabling conditions for the K-manipulating processes. The ASPI group undertook the task for creating awareness about KM within ABB, especially publicizing ABB-2 within the organization, which was a K-enabling process for ABB.

Unfortunately, ABB-1 was a fairly ineffective K-thing. We attribute this to the following problems:

1. The review process for the experiences was too cumbersome. To submit an experience, an employee needed to first submit an abstract to a review board and, if approved, send the actual experience formatted according to a detailed packaging template. The experience packages then underwent a second review before final acceptance and publishing. Turnaround times of a month or more were not uncommon for publishing an experience, exemplifying an unsuccessful K-manipulating process.
2. The database was not frequently updated because no one was formally responsible for it, further increasing the turnaround time and exemplifying an unsuccessful K-manipulating process.
3. There was no incentive to submit experiences, exemplifying lack of K-enabling process.

4. The experiences were approved only if they were a corporate-level best practice, limiting the number of experiences in the KB, and exemplifying restrictive K-manipulating policy.
5. The experience base was not well publicized, and was subsequently an unsuccessful K-enabling process.
6. It was not easy to find relevant articles in the database because the structure was not intuitive, exemplifying ineffective technology)
7. The database was limited to ABB experiences; there were no external references to supporting material, exemplifying restrictive K-manipulating policy.

In summary, the Abou-Zeid function and resources layers of the initial system (ABB-1) were lacking, reducing the utility of the system.

5.2. ABB-2: Setting quantifiable goals

At the beginning of our ABB case study, ABB-1 was the KM precedent within ABB. ABB-1 had several deficiencies in the context of the Abou-Zeid model, as listed in Section 4.1. Based on these deficiencies, the following quantifiable goals (step 2 of QIP) were identified for ABB-2:

1. make the review process simple and efficient;
2. produce an environment that encourages employees to contribute to ABB-2;
3. produce an environment that encourages employees to use ABB-2 as a source of knowledge; and
4. keep the resources updated and current.

Before designing and implementing ABB-2, it was necessary to research different types of resources that are commonly used within ABB during product or software development. After consulting the ASPI group, it was decided that the common internal resources used by BAUs for product development are templates, procedures, guidelines, and training materials. Additionally, archiving the experiences of the employees would help the BAUs share best practices. Finally, external links would enable the users to obtain easy access to resources on the Internet.

Four kinds of templates are used within ABB for product development or process improvement. These templates are differentiated based on their size and are known as:

1. Form Templates that are 1-2 pages long;
2. Document Templates that are 3-5 pages long;
3. Procedure Templates that are 6-9 pages long; and
4. Plan Templates that are greater than 10 pages.

Procedures and guidelines also exist within the organization. Training materials such as handouts, documents, and presentation slides were developed, either in-house (members of the BAUs with expertise in a particular area train other employees in the same unit) or bought from some other BAUs of ABB that possesses the required expertise. These training materials were also made available on the re-engineered KB. The experiences submitted by the employees were stored in ABB-1. These were ported to the re-engineered KB.

5.3. Choosing the process

This section describes the third QIP step, which entails determining and tailoring a process model for the project based on the context and environment of the organization. In this section, we decided the various methods and techniques to satisfy each goal identified in Section 5.2. This step produced the roadmap for developing ABB-2. The goals laid out for the deployment of ABB-2 were given in Section 5.2. The processes required to fulfill each goal is described in Table 5.1:

Table 5.1: Goal and Processes

Goal	Process(es)
Make review process simple and efficient	• Set up a review board for reviewing resources to be published on ABB-2
Produce an environment that encourages employees to contribute to ABB-2.	• Provide incentives to submit resources to ABB-2 by recognizing the top contributors • Allow a broad spectrum of resources to be published on ABB-2.
Produce an environment that encourages employees to use ABB-2 as a source of knowledge.	• Make the KB user friendly, with an intuitive user interface facilitating easy access of ABB-2 • Publicize ABB-2 within the organization so that more people know about its existence within the organization • Enrich ABB-2 with the resources already possessed by the BAUs.
Keep the resources current and updated	• Assign formal responsibility of ABB-

Goal	Process(es)
	2 to a person/group that will ensure that the KB is kept updated with new information.

ABB-2 can be visualized using the Abou-Zeid KMRM customized for ABB [39], as shown in Figure 5.1. The figure shows the three layers of the Abou-Zeid model ABB specific elements that fit into various parts of the model. In each box, the actual element of the Abou-Zeid model is given in parentheses. As seen in the figure, the B-thing for ABB is efficient and cost effective product development, and ABB-2 with artifacts such as templates, procedures, guidelines and training materials are identified as the K-thing. The ASPI team assumes the role of the actor that assumes the responsibility of the KB creator and administrator. The ASPI posts the contributions from the BAUs to ABB-2 (K-manipulating process) and supports K-enabling activities such as spreading KM awareness within ABB. Finally, the ABB Intranet and the resources of ASPI and BAUs form the resources required to develop and deploy ABB-2.

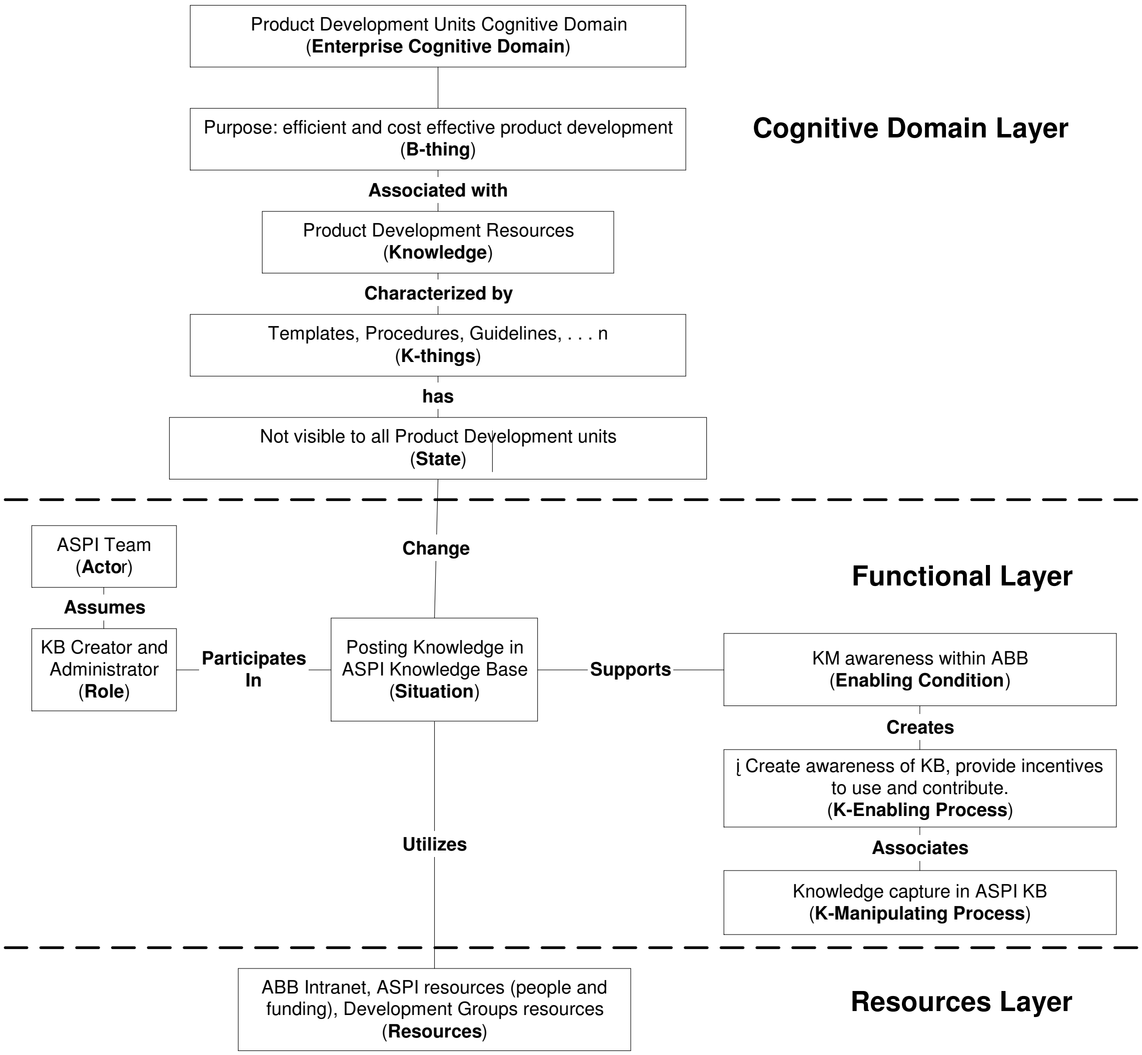

Figure 5.1: ABB KM Model Using Abou-Zeid KMRM

(Obtained with permission of authors of [39])

Thus, the Abou-Zeid model helps viewing the KM activity of deploying a KB holistically and provides a clear mapping between business objects, knowledge nuggets and resources.

5.4. Execution phase - Re-engineering the KB

This section describes ABB-2 that was developed by the author of this thesis along with a team of five members of the ASPI group. Out of these, two were in USA, two in

Switzerland and one in Sweden. In USA there was a local project leader and a principal consulting software engineer; in Switzerland, there was a local project leader and a developer, and the ASPI project leader in Sweden was mainly responsible for decision-making. The planning, design and implementation of ABB-2 was based on the goals and processes previously defined in sections 4.2 and 4.3.

5.4.1. Technology used for development

ABB-2 contains process templates, procedures, guidelines, training materials, and documented experiences of the employees in development units. Additionally, external references to non-ABB resources have been provided. More importantly, the supporting technology has been enhanced and processes have been developed to improve the utility of the system to ABB employees. ABB-2 was implemented in the Lotus Notes/Domino environment on a Windows platform. ABB-2 is a Lotus Notes database with a template-driven Web interface for easy access to the knowledge artifacts within the database. ABB-2 was developed using the Common ABB Web Platform (CAWP), an application owned by the ABB Corporate Communication division. CAWP was also used to publish the web pages on the ABB intranet. Hence, the KB has the same interface as the ABB intranet, providing visual consistency. The contents of the KB are available to all ABB employees via a well-known link on the company intranet (http://inside.abb.com/gatemodel). This link leads to the product development site of ABB that was known within ABB before ABB-2 was deployed, which helped in disseminating the contents of the KB more quickly to the end users.

5.4.2. Structure of ABB-2

ABB-2 consists of product development information for various ABB BAUs. Change agents, such as developers and project managers, can retrieve these valuable resources from the KB. In addition, employees involved in Quality Assurance and Configuration Management were also expected to use the KB. Based on the estimates from an ASPI member, ABB-2's target population was at least 1000 to 1500 employees within the organization. Any employee of ABB can submit contributions to ABB-2.

The structure of the KB ties most entries to two important reference points for the company: the ABB Gate Model [37, 38] and the SEI Capability Maturity Model Integration (CMMI) process areas [40]. Each of the CMMI process areas are related to one or more gates as defined by the ABB Gate Model. Process areas were categorized according to the Gate Model because the Gate Model was developed within ABB and is well known to all the BAUs within ABB. Since the KB entries are linked to these well-known references, ABB-2 is expected to be easy and intuitive to navigate for the users.

Figure 5.2 provides a screenshot of the home page (Level 1 page) of ABB-2. The left menu contains links to various process areas within the KB, such as requirements management. These process areas include the CMMI process areas [40]. The main page displays the gates in the ABB Gate Model as diamonds that link to resources that support each gate. The process areas underneath a gate are relevant to that gate. For example, requirements development is a part of Gate 0, Gate 1 and Gate 2. Each of these process area links leads to a process area page (Level Two page) that contains links to relevant resources. ABB-2 provides an effective capability to query for resources based up search keywords. The users can provide their feedback and ask questions through a feedback link on each page that enables the users to send an email to the administrator who is responsible for ABB-2. The top contributors to the KB are listed on the right side.

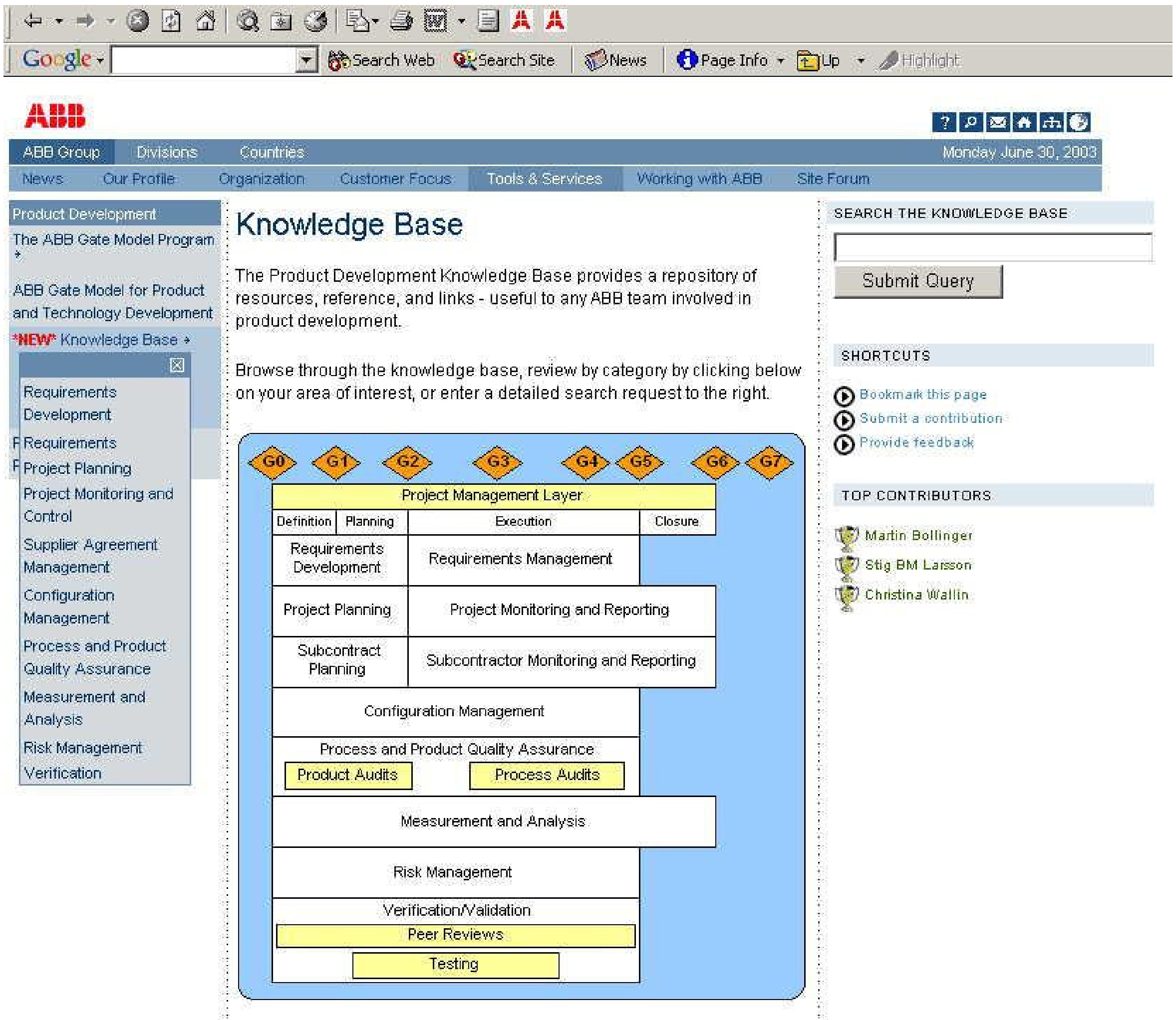

Figure 5.2: Homepage of ASPI KB

Figure 5.3 shows a sample process area page (Level Two Page). The figure shows the Configuration Management process area page. A typical process area page in ABB-2 consists of links to the following types of resources:

1. Experiences submitted by ABB employees;
2. Templates;

3. Procedures and guidelines;

4. Training materials; and

5. External references.

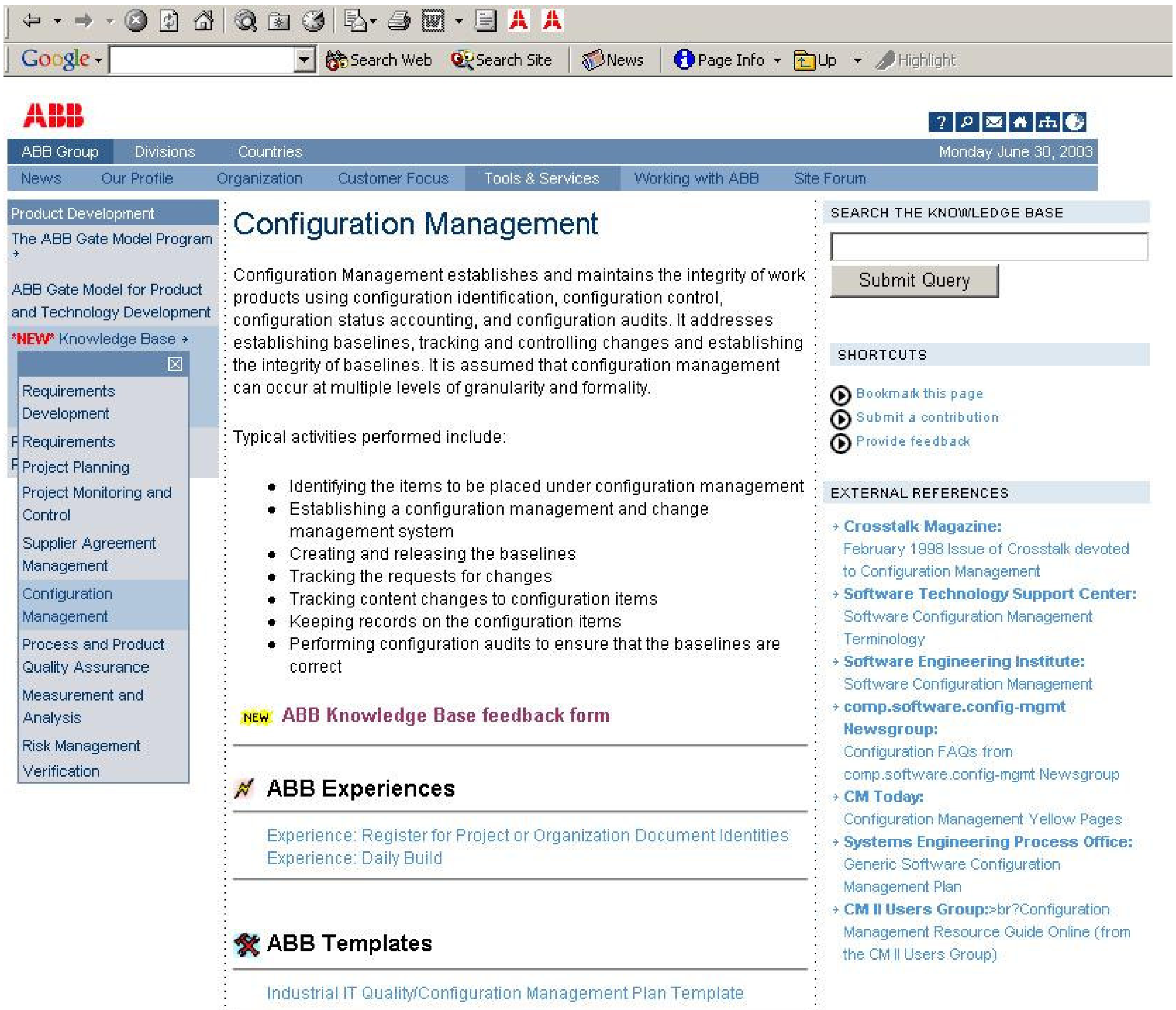

Figure 5.3: ABB-2 Process Area Page on Configuration Management

Each link leads to the corresponding resource page (Level Three page) from which the resource can be downloaded or viewed. Some resource pages may be reached from multiple process area pages. The resources themselves may have different types, such as word documents, Excel spreadsheets, or printable document files (PDF). Each resource page may also include a poll question. The poll question can be periodically changed and is used to collect feedback about the KB on some particular issue, such as ease of use or rating a particular resource. The user is offered the capability to search the KB for any specific resource or keyword through a search link. Finally, the users can provide feedback through a link that enables the user to send an email to the administrator.

The design and implementation of ABB-2 began in the last week of May 2002, and completed by October 2002. After this, ABB-2 was under review for six months and approved for publishing on the intranet in mid-March 2003. ABB-2 was published on the ABB intranet for access to all employers in early April 2003. There are currently 16 resource pages on ABB-2 that link to 13 templates, 13 experiences, and two training packages. The number of resources on the KB will increase over time as resources solicited from the different BAUs are obtained. The KB currently covers the following process areas of CMMI:

1. Requirements Management
2. Requirements Development
3. Subcontract Management
4. Project Planning
5. Project Monitoring and Control
6. Risk Management
7. Verification and Validation

Making ABB-2 user friendly was important for encouraging employees to use the KB. The two aspects of user-friendliness considered were (1) minimizing user effort to access and contribute to ABB-2 and (2) ease of use of ABB-2 through proper user interface design. Accessing the resources on ABB-2 should be easy because the KB can be accessed through well-known links within the organization. Also, to encourage contributions from the users, it was important that the mechanism to contribute be as simple as possible. Contributions to the KB and questions can be submitted through a feedback link that enables the user to e-mail the administrator. This feedback link is present on the homepage and process area pages of ABB-1.

The user interface design has been modified several times in response to feedback received from the ASPI group; this enabled us to ensure that the KB was easy to access and intuitive in the way it was organized. Another key feature was the provision of an employee incentive program to encourage contributions to this KB; a list of top contributors to provide credit and recognition is posted and regularly updated on the front page of the KB.

The contents of ABB-2 are updated regularly and new experiences submitted by the employees are incorporated after an efficient yet careful review process. The review board formally responsible for ABB-2 presently contains only one ASPI group member. The advantage of this is that it enables a swift and efficient review. The approved submissions are posted on ABB-2 within a week. However, there is a disadvantage that the contributions may be beyond the person's experience in which case he might have to refer to some one else in ABB for reviewing. This thesis recommends a few more members to take part in the review process.

5.5. Support for data collection from ABB-2

To collect data for evaluation of ABB-2, an online survey (S) and a poll (P) were created. A link has been provided in each of the process area pages (Level Two pages) to the survey S, and the poll P is a part of each resource page (Level Three pages). The online survey is accessible from each process area page as a link to a form that contains the survey questions. The survey questions can be found in the Appendix A. The survey was created using the Response Form Database, a Lotus Notes database designed for deploying online surveys for the web pages of ABB Intranet.

5.6. Future work for ABB-2

Although ABB-2 is fully functional on the intranet of ABB, it is still in its preliminary stages. ABB-2 needs to be publicized more and populated with more internal resources for stable data collection. The stages of evolution of ABB-2 are predicted as follows:

1. ABB-2 prototype implemented and functional within ABB;
2. initial publicity of ABB-2 for user response;
3. improving ABB-2 based on user feedback; and
4. begin data collection when most potential users (approx. 1000) are actively using ABB-2.

The BAUs have been requested to contribute their internal artifacts that they use for their development purposes, but initial contributions have been slow. The delay in contributions can be attributed to the restructuring of BAUs within ABB, in order to cut costs and reduce head count.

Specifically, software development process templates and training materials obtained from the BAUs will be available in ABB-2. To date, ABB-2 has been publicized informally through managers of development units. The formal publicity of ABB-2 will inform all ABB employees about ABB-2 through the ABB newsletter. Another suggested method to publicize ABB was through pop-up windows. However, the idea was not approved because pop-ups tend to annoy the users. With greater awareness of ABB-2 within the organization, it is expected that the number of users for ABB-2 will increase, and employees will regularly contribute to the KB.

6. Results

The design and implementation of ABB-2 began in the last week of May 2002 and was completed by October 2002. After this, ABB-2 was under review by members of the ASPI group for six months and was approved to be published on the intranet in mid-March 2003. ABB-2 was published on the ABB intranet for access to all employers in early April 2003. The metrics collection began in the 17th week of 2003, or the third week of April.

The rest of this chapter is organized as follows. Section 6.1 describes the results of processes executed for achieving the goals outlined in Section 5.2. Section 6.2 describes the difficulties faced in collecting data within ABB. Section 6.3 describes the data collection and data analysis for ABB. Section 6.4 describes the subjective feedback obtained from the employees through a survey and from managers through personal interviews. Finally, Section 6.5 summarizes the experiences from the ABB case study.

6.1. Analysis of goals for ABB-2

Table 6.1 describes the actual result for each process identified in Section 4.3 to satisfy the goals of Section 5.2.

Table 6.1: Analysis of Goals for ABB-2

Goal	Process	Result
Make review process simple and efficient	Set up a review board for reviewing resources to be published on ABB-2	One member of the ASPI group is currently responsible for reviewing and approving resources to be published on ABB-2.

Goal	Process	Result
Produce an environment that encourages employees to contribute to ABB-2	Provide incentives to submit resources to ABB-2 by recognizing the top contributors	Names of top contributors appear on the home page of ABB-2.
Produce an environment that encourages employees to contribute to ABB-2	Allow a broad spectrum of resources to be published on ABB-2	All types of contributions including templates, guidelines, procedures, training materials and experiences are considered. In addition to these, external references are published with each resource page.
Produce an environment that encourages employees to use ABB-2 as a source of knowledge	Make the KB user friendly, with an intuitive user interface facilitating easy access of ABB-2	The KB is structured according to process areas of CMMI and ABB Gate Model. Most of the BAUs are aware of CMMI and Gate Model. Relevant external references are placed on every resource page. The user interface is the same as ABB intranet.
Produce an environment that encourages employees	Publicize ABB-2 within the organization to facilitate its use	The publicity of ABB-2 has been delayed and still in progress. The delay was

Goal	Process	Result
to use ABB-2 as a source of knowledge		caused by top management approval of ABB-2.
Produce an environment that encourages employees to use ABB-2 as a source of knowledge	Enrich ABB-2 with the resources already possessed by the BAUs	The BAUs in direct contact with ASPI group have been solicited for resources, and four BAUs have agreed to contribute their templates.
Keep the resources current and updated	Assign formal responsibility of ABB-2 to a person/group that will ensure that the KB is kept updated with new information	One member of the ASPI group is formally responsible for maintaining ABB-2.

The use of QIP helped in identifying quantifiable goals and tracking the processes executed to satisfy the goals.

6.2. Difficulties faced in implementing ABB-2 and in collecting research data

An important goal for ABB-2 in its QIP analysis was to produce an environment that encourages employees to use ABB-2 as a knowledge source. After ABB-2 was implemented, two processes, as discussed in Section 5.2, were critical to achieve this goal:

1. The publicity of ABB-2 within the organization so that the majority of the relevant people know about its existence within the organization; and
2. The enrichment of ABB-2 with the resources already possessed by the BAUs.

ABB is a large organization and cost of communication is high. A considerable amount of time was spent in reviewing ABB-2's contents and in identifying the best location for it on the ABB intranet. This delay is attributed to the following reasons:

1. ABB recognizes the need for KM, as evidenced by the funding of ABB-2. However, significant corporate restructuring and reduction in head count took precedence over the approval of ABB-2 during the later part of 2002 and early 2003; and
2. The ASPI group has some members in Switzerland and some in USA, with the main decision making powers in Switzerland. Any decision regarding ABB-2 had to first be approved within ASPI members in USA. Final approval was then obtained from the members in Switzerland.

Any BAU requiring web statistics must pay to obtain web statistics for their site. Therefore, the web statistics required additional financial resources; approval of these funds caused further delay. This delay was worsened because of technical problems experienced when adding ABB-2 to the set of intranet pages for which statistics are collected. The web statistics were collected using the WebTrends[3] web analysis tool.

Due to the limitations of the tool and the directory structure of the intranet, some statistics collected by the tool were for the entire product development site, not specifically for ABB-2. These statistics had to be interpolated for ABB-2. Many users accessed ABB-2 through proxy servers or gateways. Hence, obtaining specific user profiles was not possible, because many of the IP addresses logged were those of proxy servers.

The web statistics show that the number of hits on the KB is low, as is expected in the early phases of a new KB. We further attribute this low activity to the lack of publicity, which hampered our ability to collect sufficient data to fully validate the metrics suite proposed in Chapter 4.

[3] http://www.webtrends.com

6.3. Analysis of case-study metrics

In this section, we describe the partial validation of the metrics suite proposed in Chapter 4. The meaning of each internal attribute measured for a KB and methods of collecting data have been discussed in Chapter 4. Data collection mechanisms have been set up for measuring each of the four internal attributes using a subset of the proposed metrics. Complete internal validation is an evolving process that requires consensus on meaning of each attribute. For example usage and effectiveness, as defined in this thesis could be defined in a different manner in a different study. Initial data collection allows partial external validation for the set of metrics. This thesis provides an analysis of the initial data collected. After substantial data is collected, the metrics can be externally validated by testing against the external attributes profit and awareness.

ABB-2 has been online for three months and metrics have been collected for ten weeks. Except for the number of hits, we present metrics for the entire ten-week period rather than give a week-by-week distribution. The metrics were collected from 3rd week of April to the last week of June.

Because ABB-2 is a part of the product development site on the ABB intranet, some of the metrics such as unique number of users accessing ABB-2 and the demographic data of users can only be collected for the entire product development site. Therefore, we can only employ aggregate metrics for the product development site as indicators of the metrics for ABB-2 when ABB-2 could not be isolated.

The other data collection methods entailed (1) using a closed-ended poll question, *P*, on the ABB-2 resource pages; (2) administering a survey, *S*, after the users had used ABB-2 for a considerable amount of time; and (3) conducting interviews, *I*, with development unit managers to obtain their feedback. The survey *S* can be found in the Appendix A. The responses to interview *I* with managers were recorded in a tabular format and are found in Appendix B. Finally, the metrics collected for ABB-2 are discussed in the following sections.

6.3.1. Usage

The following metrics have been obtained with the help of the web analysis statistics:

1. number of hits on ABB-2;
2. number of unique individuals accessing ABB-2;
3. contents of ABB-2 that are most commonly accessed; and
4. top countries using ABB-2.

A hit is a unit of user activity on the KB. If a single user browses through the same page more than once, additional visits are not included in the number of hits; instead, these additional hits are added to the number of views. ABB-2 had, on average, 32 hits per week and 64 views per week, over the ten-week measurement period. The weekly breakdown of number of hits and views for the 10 weeks is given in Appendix C. The number of hits on ABB-2 does not show an increasing trend and we believe this is because of the lack of proper publicity.

The entire product development site had 1,553 total hits during the 10 weeks for which metrics were collected. Of these, 318 hits were to ABB-2. Hence, ABB-2 accounted for approximately 20% of the hits upon being added to the already well-known ABB intranet product development site. The product development site was visited by 789 unique IP addresses, many reflecting IP addresses of proxy servers. If a computer is shared or behind a proxy server, more than one user could be using the same IP address. Of those

789 visits by unique IP addresses, 204 were repeat visitors. It is estimated that ABB-2 has 20% of the 789 or approximately 150 unique visitors and approximately 40 repeat users. Since unique visitors are identified by their IP addresses, the values of the metrics are considered a lower limit on the number of unique individuals accessing the product development site.

In cases where the IP address is traceable to an individual within the organization, the ABB's privacy policy should inform the employees that their IP addresses might be logged to collect usage metrics. ABB currently does not have a website privacy policy, and this thesis strongly recommends that ABB create a privacy policy to comply with the Fair Information Practice Principles [41]. A privacy policy is especially important given that web statistics are collected for many other purposes within ABB. A statement such as the following privacy statement about IP logging, adapted from [42], should be present in the privacy policy of the company intranet:

> *An IP address is a general representation of where you are coming from by domain name or number (i.e. aol.com, yourfirm.com, or a numeric equivalent like 206.27.160.244). We collect aggregate information site-wide, including site statistics, by logging IP addresses and/or domain names. Most often, a user's session will be tracked, but the user will remain anonymous. In some cases, the user session may be used to access the profile of the user solely for obtaining data for analysis. You may decide to "opt out" of any data analysis involving your personal profile by sending an e-mail to the website administrator.*

The five most accessed resources on ABB-2 are listed in Table 6.2. The table indicates that all kinds of resources on ABB-2 are being accessed (experiences, documents, and templates). The top-ranking resource in the table below is accessible from all the gate model links from the homepage, making it more "reachable." This explains the gap between the numbers of downloads from the most accessed and second most accessed resource in Table 6.2

Table 6.2: Most Accessed Resources on ABB-2

Resource on ABB-2 (All URL's are internal to ABB's intranet.)	Number of Downloads in second quarter of 2003
ABB Gate Model Project Management Layer	88
Experience: Dealing With A (Near) Project Failure	24
Industrial IT Pilot Requirements Specification Template	16
Industrial IT Quality/Configuration Management Plan Template	14
Industrial IT Risk Sheet	13

ABB has development units in more than 100 countries. The top countries accessing the product development site and number of hits from each country is given in Table 6.3.

Table 6.3. Top Countries Accessing ABB-2

Country of Origin for hits	Number of Hits in second quarter of 2003	Number of BAUs in the country
Sweden	419	9
Switzerland	286	8
Finland	160	3
Germany	155	7

Country of Origin for hits	Number of Hits in second quarter of 2003	Number of BAUs in the country
United States (USA)	137	14

The number of users from Sweden is greater than the number of users from the other four countries listed in Table 6.3. Since Sweden and the USA are the owners of the Gate Model, it was expected that the greatest hits for these two resources would come from the these two countries. Although Germany and Finland have the least number of development units in contact with ASPI, a good number of responses were obtained from these regions. USA is ranked fifth in Table 6.3, indicating that the product development site is not very popular in the ABB development units in the USA. ABB-2 was implemented by the ASPI group that is distributed between Switzerland and USA. Although it was expected that USA would produce the greatest number of hits, the results indicate that ABB-2 should be better publicized in the development units of ABB in USA.

The table in Appendix C indicates that although ABB-2 is operational, it is being used by a small number of employees within ABB and that the number of hits has stayed relatively constant over the 10 weeks.

6.3.2. Usefulness

A poll was administered on ABB-2. The poll question for usefulness was one closed ended question:

I was able to find the resource I was looking for in the Knowledge Base.

1. *Strongly Agree*
2. *Agree*

3. *Neutral*
4. *Disagree*
5. *Strongly Disagree*

The poll question is present on every Level Three resource page on ABB-2. Users can answer the poll question each time they access a resource. Only one employee responded to the poll, and chose the option "Neutral". This seems to be a poor response considering that there were more than 150 downloads from the KB, and the poll was present on each of the pages from which the resource was downloaded. The poll had a limitation that it was present on the same page as the actual resource, which means that the user could close down the window containing the poll after downloading the resource. This could be one of the reasons for the poor response. Using a pop-up window to attract the user's attention after the download has begun or e-mailing the poll question to potential users is of ABB-2 may be more effective than having the poll question as a passive element of the resource page.

The average time spent on each page of the KB was collected via the WebTrends web analysis tool. The results show that the ABB-2 home page was viewed for 33 seconds per hit on average. Although this seems a very short amount of time, it is important to consider that the home page (shown in Figure 5.2) consists only of links to process area pages rather than any descriptive material or meaningful content. The user is expected to click on one of the process areas and promptly leave the home page. Also, a repeat user would spend even less time on the home page because he or she might know where to click to go to a Level Two page.

The top five Level Two pages on which the users spent most time, in decreasing order, are shown in Table 6.4. Again, all these times are less than a minute because Level Two pages also contain links to various types of resources. Typically, a user rapidly clicks on a link to one of the resources and proceeds to the Level Three page to access the resource.

Table 6.4: Level Two Pages with Greatest Time Spent

Level Two page (All URL's are internal to ABB's Intranet)	**Average time (seconds)**
Personal Software Process(PSP)/Team Software Process (TSP)	49
Requirements Development	46
Project Planning	41
Risk Management	36
Project Monitoring and Control	29

Level Three Pages are the resource pages that have direct link to the actual resource. The users browsing on resource pages may either view the resource online, or download the resource. Hence, the time spent on resource pages is not consistent for comparison.

We now describe the subjective feedback obtained from the survey *S* for usefulness. Five users responded to the survey *S*. Four of the responses were obtained through an email sent directly to the employees by a member of the ASPI group. This member had the e-mail addresses of a set of individuals that was aware of ABB-2, and therefore might be more willing to take the survey. One person voluntarily took the online survey. Out of the five responses, there were two each from Switzerland and Sweden and one from Germany.

No one from the USA took the survey, another indication that greater publicity is needed in the USA. The following results were obtained with respect to usefulness:

- Four out of five respondents agreed that ABB-2 made it easier to find product development information. One respondent was neutral.
- Two respondents frequently used the KB (at least once a week) while the other three did not use it that frequently.

Summarizing, the data obtained so far is not enough to draw any conclusions. However, if these results were based on a large response, then it would an indication that most of the people find ABB-2 are finding the ABB-2 useful. However, ABB-2 is not frequently used by more than half of the users, which leads to many possibilities, such as:

1. Since the BAUs already have most of the artifacts, the users do not need to access ABB-2 that often.
2. Accessing information from ABB-2 is still cumbersome, deterring users from taking the effort to find artifacts on ABB-2;
3. ABB-2 needs to be better populated with resources; or
4. Users may be downloading many resources at a time, and hence need to visit ABB-2 less often.

As more people respond to the survey that is still online, more substantial results may be obtained and a more conclusive analysis can be done. The poll *P* is online and more responses are expected when the number of users of ABB-2 increases.

6.3.3. Effectiveness

The following survey question regarding effectiveness was a part of survey *S:*

What percentage of times does the KB produce useful information?

All the five respondents believe that the KB provides useful information less than 60 percent of the time. Had this data been obtained from a large population, it would indicate the following possibilities:

1. ABB-2 may need to be populated with more resources;
2. ABB-2 may not be targeting the correct set of users;
3. ABB-2 structure needs to be reconsidered because users may not be able to intuitively obtain the required resource, although it is present on ABB-2.

Each of these possibilities would have to be further explored through user feedback, and corrective action taken.

The frequency of contribution metric was measured via the following question in survey S.

How many articles/experiences do you intend to contribute to the Knowledge Base in the next six months?

Based on the results from the five responses, two contributions per survey respondent can be expected within the next six months.

Initial data shows that ABB-2 is not very effective in providing useful product development resources to the users. Considering that ABB-2 is still in its early days of usage, publicity of ABB-2 and contributions from the BAUs are two critical aspects for the success of the ASPI KB.

6.3.4. Return on Investment

The objective of ROI metrics was to determine whether the benefits from ABB-2 outweigh the effort involved in designing, implementing, deploying, and maintaining the KB. Specifically, we quantify ROI by the time saved within the organization because of

ABB-2. Before ABB-2, there was no common information repository to product development resources within ABB. ABB-1 only provided experiences and training packages and did not include templates, procedures, and guidelines that are often used. For example, many versions of the same template existed with different development units within the organization. In the absence of ABB-2, many of the already-existing templates were re-developed from scratch were based upon other web resources (such as IEEE standards).

ABB-2 is intended to minimize re-development of artifacts, and the BAUs have been solicited to provide their templates to the KB. As soon as the templates are received by the ASPI member responsible for ABB-2, they will be categorized and similar templates (e.g. from the process area of configuration management) will be consolidated. Only a few variations of the each template type will be published. Other BAUs that do not possess a template can download a version from the KB and customize it, if necessary, to their specific purpose. ABB-2 also contains training materials. The reuse of training material throughout the organization is a factor for timesaving because of the presence of ABB-2.

Our analysis is inspired by the fact that reuse within the development units due to ABB-2 can be quantified by finding out how many artifacts are reused and using the approximate time saved due to reuse of each artifact to find the time saved within the organization. This information can provided through interviews with the managers of BAUs who have considerable experience in overseeing development projects. Through their experience, the managers can provide a fair estimate of time saved due to reuse on each artifact, and the number of artifacts of each type developed using the KB.

For our ROI analysis, a representative sample of four managers from different BAUs in the USA were each interviewed (interview I) to collect information about the number of artifacts present in the BAU and the time spent to develop and customize these artifacts. Each of the managers interviewed had at least ten years of experience with one or more BAUs in ABB. The fifth interview was conducted with an ASPI group member who had 12 years of experience in developing templates and training materials for different BAUs within ABB. The author of this thesis conducted all five interviews, after ABB-2 had been online for three months. Each interview was about 15 minutes long. Two of the interviews were telephonic interviews and three interviews were face-to-face interviews.

The interviews specifically had questions about templates and training materials. The managers and the ASPI group member were informed about the intent of the interview and were told that their responses shall be kept anonymous in the thesis. The following questions were asked for each type of template listed in Section 5.2:

1. On average, how much time is spent in developing a template from scratch?
2. On average, how much time is spent in customizing a template?
3. How many templates of each type will a particular business unit contribute to ABB-2?
4. How many documented templates are present and how many templates have to be documented in the near future?

The following questions were asked regarding training materials:

1. On average, how much time is spent in preparing training material for a one-hour training session?
2. On average, how much time is spent in customizing training material for a one-hour training session?

3. How many hours of training per month is given to the employees in your business unit?
4. How many training packages will you contribute to ABB-2?

Additionally, the experience of the managers in management and development within their BAUs (their career profiles) was noted, ensuring that the information obtained from them is reasonably credible. The managers also provided input on what they would like from a KB and commented on the already existing ABB-2.

The managers of development units were asked to provide the approximate time required for developing and customizing an artifact. For every type of template described in Section 5.2, the difference between development and customization time was calculated, using the values recorded from interviews. This difference was then multiplied by the average number of each type of template in a BAU to calculate the total time saved for each BAU due to templates. A similar calculation was made for training material. The total time saved for each BAU was then multiplied by the total number of BAUs accessing ABB-2 to find approximate time saved in the entire organization. We now provide mathematical representations for the method just described.

Let NetTime(*artifact*) be the difference between the development time (NetTime(*artifact*)) and customization time (customizeTime(*artifact*)) for a particular artifact. This is summarized by equation E-1:

E-1: NetTime(*artifact*) = developTime(*artifact*) – customizeTime(*artifact*)

Equation E-1 is generic, in the sense that *artifact* can represent any type of template or training material. For example, the time saved for each form template would be:

NetTime(Form Template) = developTime(Form Template) – customizeTime(Form Template)

Typical values of developTime and customizeTime for each artifact were obtained from each response from the interview I. NetTime(*artifact*) was calculated for each type of artifact, for each response. Table 6.5 shows the average value of NetTime(*artifact*) for each type of artifact based on our small sample size of interviewees.

Table 6.5: Average Time Saved per Artifact

Type of artifact	**Person-hours saved** **i.e. avg(NetTime(*artifact*))**
Form Template	5.4
Document Template	15.5
Procedure Template	25.7
Plan Template	37.5
Training material (per hour of training)	10.3

Templates and training materials are now treated separately. Let N(*template*) be the number of template of a particular type. Separate calculations were made for templates already developed and for the templates to be developed in the near future (based on the responses from the managers). Let N_1(*template*) represent the actual number of developed templates and N_2(*template*) represent the projected number of templates to be developed in the future.

Table 6.6 shows the average number of documented and undocumented templates. These are calculated by averaging N_1(*template*) and N_2(*template*) based upon the responses of the five interviewees. Three of the five respondents said they currently had all templates documented and would need to document more templates only if they went for certain ISO certifications. But they could not predict when they will need to document more templates. This thesis assumes zero undocumented templates for these three. Hence, this thesis makes a conservative estimate concerning the undocumented templates, which reduces the projected timesaving.

Table 6.6: Average Number of Templates per BAU

Type of template	**Documented Templates ($avg(N_1(\textit{template}))$)**	**Undocumented templates ($avg(N_2(\textit{template}))$)**
Form Template	20.0	2.2
Document Template	6.2	1.2
Procedure Template	14.2	1.8
Plan Template	2.8	0.2

Approximate savings for each type of template is then calculated through the generic equation:

E-2: TimeSaved(*template*) = avg(NetTime(*template*)) * avg(N(*template*))

Let $TimeSaved_1$ and $TimeSaved_2$ denote the time saved from documented and undocumented templates per BAU, respectively. For example, the projected timesaving for undeveloped form templates is:

$TimeSaved_2$(Form Template) = avg(NetTime(Form Template)) * avg(N_2(Form Template))

Table 6.7 lists the values of $TimeSaved_1$ and $TimeSaved_2$ for each type of template.

Table 6.7: Projected Timesaving (Hours) per BAU because of Templates

Type of template	**Retrospective time savings** ($TimeSaved_1$(*template*))	**Projected time savings** ($TimeSaved_2$(*template*))
Form Template	107.5	11.8
Document Template	96.1	18.6
Procedure Template	364.5	46.2
Plan Template	105.0	7.5

Note that the templates already developed were developed without the aid of archived templates on ABB-2. Mathematically, the number of person-hours that could have been saved in each BAU had the development been done using a KB can be written as equation E-3:

E-3: $RetrospectiveSaving = \sum_{(\text{For each type of Template})} TimeSaved_1(template)$

Similarly, the projected number of person-hours that will be saved in the future because of the presence of ABB-2 is written as equation E-4:

E-4: $ProjectedSaving = \sum_{(\text{For each type of Template})} TimeSaved_2(template)$

Ideally, the time saved should be calculated by determining the actual number of templates developed using resources from ABB-2 and then multiplied by the difference between development and customization times. However, the fact that ABB-2 has not yet been publicized properly, limiting its use within the organization, called for a futuristic evaluation in this case study. The actual number of templates developed using resources from ABB-2 will keep changing from time to time, which means that the calculation for the time saved within the organization, is temporal and will be valid only for a particular time period.

Using E-3 and E-4, RetrospectiveSaving = 673.1 person-hours and ProjectedSaving = 84.1 person-hours. Although ABB has around 100 BAUs that concentrate on product development, only 41 BAUs use the services of the ASPI group. Although the resources from ABB-2 will be available to all the BAUs of ABB, this thesis makes a conservative estimate by considering only the 41 BAUs that are in direct contact with ASPI group in our calculations. Hence, the retrospective time savings and projected time savings with respect to templates is:

Retrospective time savings for Templates = 673.1 * 41 = 27,597 person-hours

Projected time savings for Templates = 84.125 * 41 = 3,449 person-hours

The timesaving for training materials was calculated in a similar fashion, but on a per-month basis. The time saved per month per BAU because of training materials ($TimeSaved_3$) can be written as Equation E-5:

E-5: $TimeSaved_3$ = No. of hours of training using KB * NetTime(Training Material)

The value of NetTime(Training Material) is obtained using equation E-1, and the average value is 10.1 as listed in Table 6.4. The responses from interview I indicated that there is 4.2 hours of training per month per BAU or around 50 hours of training per year per BAU. Training is provided free of cost by ASPI to the other groups within the organization. Typically, one ASPI member provides training to one member of the on-site BAU (known as the change agent), and the change agent is then responsible to provide in-house training to other members of the BAU. This way, the ASPI group aims to make the BAUs as self-sufficient as possible, and ABB-2 is a supporting tool for this vision. Change agents of different BAUs require training material, and ABB-2 will facilitate finding ABB specific training material. Based on input from the ASPI group, approximately eight hours out of the 50 hours of training per year is done with the support of ASPI group for each BAU it is in direct contact with. Since the training is new most of the time for the change agents, 75% or six hours of ASPI training may be considered as training on new areas, done by change agents of the BAU. Hence, time saved per BAU per month on training is 6 * 10.1 = 120.6 person-hours/year/BAU. Therefore, total time saved in the entire organization per year is given as:

Time saved for ASPI Training = 120.6 * 41 = 4,945 person-hours

The approximate amount of money spent on ABB-2 thus far is $117,500. Of this, approximately $34,000 was paid to the author of this thesis who worked for approximately 1000 hours on ABB-2. The remaining $83,500 was allocated for the time of other ASPI members and a small amount was spent for paying for web statistics. We take a conservative estimate by ignoring the money spent on obtaining web statistics and assume the standard pay of $117/hr within ABB. Then, the ASPI members spent approximately 713 hours on the KB, leading to total of 1,713 hours already spent on ABB-2. Additionally, 400 person-hours per year will have to be dedicated for reviewing new contributions and posting it to ABB-2. These estimates have been obtained from the ASPI group. Table 6.8 shows the summary of benefits of using ABB-2:

Table 6.8: Projected Timesaving versus Investment

Total investment in ABB-2	1,713 person-hours + 400 person-hours/year
Projected time savings because of templates in the near future	3,449 person-hours
Projected time savings because of ASPI training materials	4,945 person-hours/year

The above results show that the projected timesaving from ABB-2 is much greater than the amount of time invested in ABB-2 even by using conservative estimates of the time saved. With proper publicity and constant updating, the benefits could be potentially greater than those calculated. It is also notable that these calculations do not include the benefits of experiences, external links to resources, and improved ease of contacting experts. Also, the results are temporal because these figures could change as more resources are added to ABB-2 or through interviews with other development units.

6.4. Subjective feedback

The users who took the survey S were asked to list the strengths and weaknesses of ABB-2, and what information should be available from ABB-2. Although these responses are anecdotal at best, it is interesting to note the following user responses:

> "It's not easy to find, and it's poorly marketed"

"A news-group about product development for ABB-Members, where the user can discuss and get some answers from the other users."

"I like the "EXTERNAL REFERENCES" and have already used them . . ."

ABB-2 is in its early days of adoption, due to which the data collection from the users is still incomplete. However, initial data indicates the strengths and weaknesses of ABB-2. Firstly, ABB-2 needs to be reinforced with a forum enabling users to discuss particular topics of interest. Also, references to ABB resources may not always satisfy the users' needs. In such cases, external references may provide the required resources.

Subjective feedback was also received from the development unit managers through interview I. The managers were asked what information should be available on ABB-2. Their responses are given below:

- One respondent expressed that ABB-2 should have a "Frequently Asked Questions" section about common difficulties in product development and a "Tools and Techniques" section with links to common product development tools. He also said that more resources are required regarding collecting metrics for projects.
- One manager believes the templates described by IEEE standards and the like are usually for big groups, and there should be templates customized for small groups involved in software or product development.
- One respondent said that ABB-2 should have sections specifically catering to groups that were going for ISO to CMM transition. Also, metrics for measuring software size for real-time control systems should be posted.

6.5. Lessons Learned

A KB is successful in an organization only if it benefits the intended users. The documentation of lesson learned is the packaging step (last step) for QIP. The following lessons were learned from the ABB case study:

1. A KM framework helps guiding a KM project. The framework can be used to outline the goals for the project and also to uncover potential pitfalls. For example, the Abou-Zeid reference model helped in the retrospective analysis of the deficiencies of ABB-1, which were removed in ABB-2.
2. One member of the ASPI group was assigned formal responsibility of maintaining and updating ABB-2. This was important because no one was formally responsible for maintaining ABB-1, causing it to be not frequently updated. After it was developed, ABB-1 remained in the organization for a long time without being used or updated. When one or more people take responsibility of a KB, they ensure that they find time to review contributions and update the KB.
3. The metrics collected for ABB-2 indicate that the KB is used by a very small number of employees within the organization. Proper publicity is very important to encourage employees to use a KB. The publicity at ABB was performed by word of mouth publicity, rather than publicize the KB through the company newsletter, or through an online campaign. Although an important factor at ABB was the restructuring and head count reduction taking place inside the organization, this thesis proposes a more formal publicity drive.
4. The responses from the users indicate that ABB-2 is not populated with a large number of resources. Therefore, the users fail to find useful information at least 40% of the

time, which might discourage them on future use of the KB. The delay in obtaining resources already available in individual BAUs hampered updating the KB. The KB should be well populated with valuable resources to significantly increase users' ability to find the desired resources and to feel the KB is worthwhile.

5. Proper pointers to external resources help the users to find additional information from the WWW. The web links to external resources are present on each Level Two page. If a user is not able to find a particular resource for a particular process area amongst the internal ABB resources, he could find them in one of the external references.
6. Personal contact for feedback worked better than soliciting feedback through passive ways. A personal e-mail from one of the ASPI group members and telephone interviews were the two successful data collection techniques. Feedback forms and polls linked from the KB did not produce a good response. The poll had just one response. It would be better to have the poll as a pop-up window or for the poll to be sent through personal e-mail.
7. Geographical disparities in the usage of ABB-2 helped in determining regions where ABB-2 was not so popular, calling for better publicity and awareness in those regions. There could be other reasons for these disparities, such as restructuring within the organization, cultural issues, or more projects requiring artifacts from the KB within one region than another.
8. The most accessed resource from ABB-2 (refer Table 6.2) has now been converted to HTML format for easy access. It may be a good idea to convert resources (wherever possible) to online HTML documents, enabling users to preview the document and

then make the choice of downloading it or not. This would make ABB-2 user-friendlier, encouraging more users to use ABB-2.

7. Conclusions and Future Work

A metric suite for the evaluation of a KB was presented in this thesis. Metrics were identified for four types of internal attributes, which can be used to assess a KB: usage, usefulness, effectiveness, and ROI. This thesis recognizes that ROI is difficult to measure. The method proposed in this thesis to calculate ROI produces an approximate comparison between the efforts required for building a KB versus the effort saved because of the presence of the KB. Once metrics are collected, it is important to make sure that the metrics chosen are the right ones to evaluate the KB. Internally validating the metrics involves checking that a metric really measures what it purports to measure [34]. External validation involves correlating the actual values of the metric with some external attribute. The two external attributes proposed by this thesis are profit to the organization due to the KB and awareness about the KB within the organization. This thesis proposes validating usage, usefulness and effectiveness metrics against both profit and awareness as the external attributes and ROI should be validated against profit as the external attribute.

One goal of this thesis was to describe the re-engineering effort at ABB. The re-engineering effort can be seen in the context of the reference model proposed by Abou-Zeid. The Abou-Zeid model helped in identifying the deficiencies in ABB-1 and in deciding the goals for ABB-2. The six steps of QIP guided the re-engineering process. Using an appropriate model for KM within a company and an appropriate process for guiding a KM project, an organization can systematically implement a KM project.

The ABB case study can serve as a guide to implementing a KB within an organization. Based on the experience with the ABB case study, this thesis acknowledges that four factors are instrumental to the success of a KB:

1. assigning formal responsibility of the maintaining the KB to one or more persons;
2. proper publicity of the KB within the organization;
3. ability to find resources on KB, making it conducive to visit the KB for knowledge needs; and
4. employee participation in contributing to the KB, and keeping the KB regularly updated.

A subset of the metrics proposed in Chapter 4 was used to evaluate ABB-2. The usage metrics show that ABB-2 is being used by only a small number of users. Although four out of five users felt that ABB-2 made it easy to find useful information, all five users believe that ABB-2 produces useful information less than 60% of the times. However, the projected timesaving because of the presence of ABB-2 is much greater than the time invested in developing and evaluating ABB-2, even by taking conservative estimates. Partial internal validation was conducted by clarifying each metric and partial external validation was performed by collecting the initial data for the metrics.

ABB-2 is being continuously populated with resources. However, to increase ABB-2's awareness, publicity must be increased, especially in countries such as USA where ABB-2 obtained a less than expected response. The data collection mechanisms are in place so that data can be collected for a more comprehensive analysis when ABB-2 is regularly used by a large number of employees (at least 1000 unique users). After data for the metrics are obtained, their association with external attributes such as profitability and awareness should be tested for complete validation.

8. List of References

[1] M. Lindvall, I. Rus, R. Jammalamadaka, and R. Thakker, "Software tools for Knowledge Management," DoD Data and Analysis Center for Software, Rome, N.Y. 2001.

[2] L. J. Bassi and M. E. V. Buren, "New Measures for a New Era," in *Knowledge Management: Classic and Contemporary Works*, B. Thuraisingham, Ed.: MIT Press, 2000, pp. 355-373.

[3] I. Rus and M. Lindvall, "Knowledge Management in Software Engineering," *IEEE Software*, vol. 19, pp. 26-38, May/June 2002.

[4] A. Tiwana, *The Knowledge Management Toolkit: Practical Techniques for Building a Knowledge Management System*: Prentice Hall PTR, 2000.

[5] I. Rus, M. Lindvall, and S. S. Sinha, "Knowledge Management in Software Engineering," DoD Data and Analysis Center for Software, Rome, NY 2001.

[6] E. Swanstrom, "Metrics: Separating KM Fact from Fiction," in *Knowledge Management: Classic and Contemporary Works*, B. Thuraisingham, Ed.: MIT Press, 2000, pp. 313-315.

[7] H. Takeuchi and I. Nonaka, *The Knowledge Creating Company:How Japanese Companies Create the Dynamics of Innovation*: New York: Oxford University Press, 1995.

[8] S. Komi-Sirvio, A. Mantyniemi, and V. Seppanen, "Toward a Practical Solution for Capturing Knowledge for Software Projects," *IEEE Software*, vol. 19, pp. 60-62, May/June 2002.

[9] T. K. A. Hamid, "The Slippery path to productivity improvement," *IEEE Software*, vol. 13, pp. 43-52, 1996.

[10] V. Basili and S. Green, "Software Process Evolution at the SEL," *IEEE Software*, vol. 11, pp. 58-66, July/August 1994.

[11] V. Basili, G. Caldiera, and H. D. Rombach, "Goal Question Metric Paradigm," in *Encyclopedia of Software Engineering*, J. J. Marciniak, Ed.: John Wiley Sons, 1994.

[12] V. Basili and H. D. Rombach, "The TAME project: Towards Improvement-Oriented Software Environments," *IEEE Transactions on Software Engineering*, vol. 14, pp. 758-773, June 1988.

[13] V. Basili and D. Weiss, "A Methodology for Collecting Valid Software Engineering Data," *IEEE Transactions on Software Engineering*, vol. 10, pp. 728-738, November 1984.

[14] V. Basili, "Software Modeling and Measurement: The Goal Question Metric Paradigm," *Computer Science Technical Report Series*, September 1992.

[15] S. Ramasubramaniam and G. Jagadeesan, "Knowledge Management at Infosys," *IEEE Software*, vol. 19, pp. 53-55, May/June 2002.

[16] C.-P. Wei, P. J.-H. Hu, and H.-H. Chen, "Design and Evaluation of a Knowledge Management System," *IEEE Software*, vol. 19, pp. 56-59, May/June 2002.

[17] V. Basili, M. Lindvall, I. Rus, C. Seaman, and B. Boehm, "Lessons-Learned Repository for COTS-Based SW Development," *Software Technology Newsletter*, vol. 5, September 2002.

[18] B. Rubenstein-Montano, J. Liebowitz, J. Buchwalter, D. McCaw, and The Knowledge Management Methodology Team, "A Systems Thinking Framework for Knowledge Management," *Decision Support Systems*, vol. 31, pp. 5-16, May 2001.

[19] D. Binney, "The Knowledge management Spectrum - Understanding the KM Landscape," *Journal of Knowledge Management*, vol. 5, pp. 33-42, 2001.

[20] K. Ehms and D. M. Langen, "Holistic Development of Knowledge Management with KMMM®," presented at American Productivity & Quality Center (APQC), Las Vegas, 2000.

[21] M. C. Paulk, B. Curtis, M. B. Chrissis, and C. V. Weber, "Capability Maturity Model for Software, Version 1.1," *IEEE Software*, vol. 10, pp. 18-27, July 1993.

[22] E.-S. Abou-Zeid, "A Knowledge Management Reference Model," *Journal of Knowledge Management*, vol. 6, pp. 486-499, 2002.

[23] V. Basili, "Quantitative evaluation of software engineering methodology," presented at First Pan Pacific Computer Conference, Melbourne, Australia, September 1985.

[24] A. Maclean, R. Young, and T. Moran, "Design Rationale: The Argument Behind the Artifact," presented at SIGCHI conference on Human factors in computing systems, May 1989.

[25] P. Louridas and P. Loucopoulos, "A Generic Model for Reflective Design," *ACM Transactions on Software Engineering and Methodology*, vol. 9, pp. 199-237, 2000.

[26] T. Moran and J. Carroll, *Design Rationale: Concepts, Techniques and Use*. Mahwah, NJ: Laurence Erlbaum Associates, Inc., 1996.

[27] L. Karsenty, "An Empirical Evaluation of Design Rationale Documents," presented at Human Factors in Computing Systems, Vancouver, Canada, 1996.

[28] G. Arango, E. Schoen, and R. Pettengill, "A Process for Consolidating and Reusing Design Knowledge," presented at International Conference on Software Engineering, Baltimore, Maryland, 1993.

[29] D. Skyrme, "Developing a Knowledge Strategy: From Management to Leadership," in *Knowledge Management: Classic and Contemporary Works*, B. Thuraisingham, Ed.: MIT Press, 2000, pp. 61-83.

[30] D. Cohen, R. Azzarello, L. Prusak, and D. Smith, *Managing Knowledge for Business Success: A Conference Report*: Conference Board Paperback, 1997.

[31] R. Camp, *Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance*. Milwaukee, WI: ASQC Quality Press, 1989.

[32] K. Lopez, "How to Measure the Value of Knowledge Management," *Knowledge Management Review*, vol. 4, pp. 20-23, 2001.

[33] W. Vestal, "Measuring Knowledge Management." http://www.kmadvantage.com/docs/km_articles/Measuring_KM.pdf, August 2002.

[34] K. E. Emam, "A Methodology for Validating Software Product Metrics," National Research Council of Canada NRC/ERB-1076, June 2000.

[35] N. E. Fenton, "Software Metrics: Theory, Tools and Validation," ***Software Engineering Journal*****, vol. 5, pp. 65-78, January 1990.**

[36] E. Taylor-Powell and M. Renner, "Analyzing Qualitative Data." http://www1.uwex.edu/ces/pubs/pdf/G3658_12.PDF, 2003.

[37] R. A. Carter, A. I. Antón, A. Dagnino, and L. Williams, "Evolving Beyond Requirements Creep: A Risk-Based Evolutionary Prototyping Model," presented at IEEE 5th International Symposium on Requirements Engineering (RE'01), Toronto, Canada, 2001.

[38] C. Wallin, F. Ekdahl, and S. Larsson, "Integrating Business and Software Development Models," vol. 19, pp. 28-33, November/December 2002.

[39] P. Baheti, L. Williams, A. Dagnino, and A. Cordes, "On Identifying Deficiencies in a Knowledge Management System," presented at Fifteenth International Conference on Software Engineering and Knowledge Engineering, San Francisco Bay, CA, 2003.

[40] M. B. Chrissis, M. Konrad, and S. Shrum, ***CMMI®: Guidelines for Process Integration and Product Improvement*****: Addison Wesley Professional, 2003.**

[41] U.S. Department of Health, Education and Welfare, and R. Secretary's Advisory Committee on Automated Personal Data Systems, Computers, and the Rights of Citizens, "The Code of Fair Information Practices," http://www.epic.org/privacy/consumer/code_fair_info.html, 1973.

[42] "Privacy Policy of ProActive Health." http://www.proactivehealth.com/PrivacyPolicy.asp.

9. Appendices

Appendix A: Survey S for ABB-2 users

1. The Knowledge Base makes it easier to find product development information.

- Strongly Agree
- Agree
- Neutral
- Disagree
- Strongly Disagree

2. How often do you refer to the Knowledge Base?

- Daily
- At least once a week
- At least once a month
- At least once in 3 months
- Less frequent than once in 3 months

3. What percentage of the times does the Knowledge Base provide useful information?

- < 20
- 20 – 39
- 40 – 59
- 60 – 79
- 80 – 100

4. What weaknesses do you find in the Knowledge Base?

5. What other information would you like to see on the Knowledge Base?

6. How many articles/experiences do you intend to contribute to the Knowledge Base in the next six months?

Appendix B: Interview I for Managers of Development Units

Part 1: Interview Sheet for Noting Responses

Name of Interviewee:

	Development Time (person-hours)	**Customization Time** (person-hours)	**Approx. Savings** (person-hours)
Form Template (1-2 pages)			
Document Template (3-5 pages)			
Procedure (6-9 pages)			
Plan (> 9 pages)			
One hour training			

	Documented	**Undocumented**	**Total**
Form Template			
Document Template			
Procedure			
Plan			
# of types of training			
Average length of training			

Contributions To KB	Form Template	Document Template	Procedure	Plan	Training Materials

Part 2: Interview Questions

1. Describe your development unit, and your experience in software/product development.

2. Do your templates filled by engineers contain instructions on how the form is to be filled?

3. What information, if present, on the ASPI Knowledge Base would save you time?

4. Any other comments?

Appendix C: Number of hits (and views) for ABB-2

Week number (2003)	Number of hits	Number of views
17	38	98
18	25	54
19	30	54
20	33	81
21	24	40
22	25	40
23	39	63
24	28	55
25	42	74
26	34	80