Chapter 2. Knowledge Management for Digital Preservation

2.1. Introduction

The major threats of long term preservation are the changes over the long term. Generally, the changes include not only the preserved knowledge or data themselves, but also the Knowledge Base inside people’s mind. Either change will reduce the traceability and reusability of appropriate knowledge. As we have discussed in the previous chapter, in order to achieve long term knowledge preservation, we need to develop a KM approach as well as a digital preservation approach. In this chapter, we develop a KM methodology for providing knowledge objects for digital preservation. The KM approach needs the strategic decision making, planning and functional processes for knowledge management and knowledge retention. Based on the preliminary research on KM methodologies, we choose to extend from an existing KM methodology (i.e. CommonKADS) in our research, in order to adapt the Product-Process-Organization (PPO) design principle in knowledge management approach for products.

2.2. Knowledge Management Methodology

As we have already discussed in the previous chapter, “knowledge retention” is in the “knowledge management” context. And we have to study KM, in order to develop approaches for “knowledge retention”. In this section we have studied KM methodology, tools and systems. The KM methodologies will help to develop methodological approaches and models for KM and thus knowledge retention. The KM tools provide functionalities and features to achieve the knowledge retention approaches and models. The KM system is a composition of KM technologies and tools, in order to support KM. In our research work, we introduce the “system” concept, in order to establish a platform for the perspective of long term digital preservation.
2.2.1. Knowledge and Knowledge Management

The formal definition of “knowledge” is defined by Webster’s dictionary and implies that knowledge extends beyond information. It gives the following description:

Knowledge - noun. 1. applies to facts or ideas acquired by study, investigation, observation, or experience  2. rich in the knowledge of human nature  3. Learning applies to knowledge acquired especially through formal, often advanced, schooling  4. a book that demonstrates vast learning.

Data, information and knowledge are three often encountered words that are close together. However, data refers to a collection of facts usually collected as the result of experience, observation or experiment, or processes within a computer system, or a set of premises. Information is the basis of knowledge, might be directly associated with the facts of the real world. And knowledge often depends on the context, thus one person’s knowledge could be another person’s information.

Knowledge Management (KM) is the practice of selectively applying knowledge from previous experiences of decision-making to current and future decision-making activities with the express purpose of improving organizational effectiveness [Zhang 04]. For enterprises and industries, an archive of knowledge must capture all of the data required to completely define the product, and in some instances, processes. Thus KM is not just a technology, but it is about people, processes and practice.

Our research concerning knowledge management represents initial work for long term digital preservation. According to the study of Dave [Dave 07], the discipline of KM has changed very quickly over the last decade, and KM can generally fall into two generations. The first generation of KM is from 1995 to 2005, and the second generation is from 2005 to today. Besides the strategic differences between the two generations of KM (from central large repositories to shared personal repositories), the content formats of the first generation of KM are mostly categories of texts organized by subjects (taxonomy), while the content formats of the second generation of KM are graphics and multimedia organized by applications (ontology). Therefore, in modern KM discipline, managing knowledge activities are performed in a collaborative environment [Sureephong 09]. The complexity of producing, keeping and reusing knowledge from graphic and multimedia formats implies that digital information preservation approaches and technologies play an important role in modern KM projects.
2.2.2. Knowledge Management Methodologies

In order to achieve dynamic preservation and keep the preserved knowledge and data always up-to-date, we need to develop the strategic decision making, planning and functional processes for knowledge management and knowledge retention. Therefore, we need to adapt the KM approach and develop a methodology.

Ritendra Banerjee in his article [Banerjee 05] claims that the KM Cycle is described as

- Knowledge Generation
- Knowledge Codification
- Knowledge Retrieval
- Knowledge Transfer
- Knowledge Purging

Organizations need to determine a formal process in order to organize knowledge: identify, capture, store, and retrieve critical knowledge. Thus, organizations need KM to help them identify what they know, what they need to know, and how to effectively use what they know [Jennex 08]. According to the generalization of KM methodologies, the following KM steps will be used in our research:

- Researching and aligning with Corporate Strategy
- Identifying business process and knowledge
- Capturing knowledge
- Communicating and organizing knowledge
- Creating a knowledge-sharing culture
- Benchmarking
- Improving the process continuously

Although via the definition of operational steps of KM methodology we get references for the deployment of long term knowledge preservation, we still need recommendation design models for the process of organizing knowledge to reach the enterprise needs of preservation, as well as maintenance and transaction of knowledge. In other words, we have to develop knowledge engineering methodology to carry out long term knowledge preservation.

Knowledge engineering refers to the designing, developing and maintaining of Knowledge Based Systems (KBS) in the knowledge management project. It has a great deal in common
with software engineering [Sureephong 09], and is related to many computer science domains such as artificial intelligence, databases, data mining, expert systems and decision support systems. Our methodology of long term knowledge preservation has established a dynamic strategy and preservation system to keep the stored knowledge (Product-Process-Organization: PPO model of PLM in enterprise [Noël 08]) always up-to-date; thus, to overcome the long term changes. Through the research on knowledge management and digital preservation, we notice that there is gap between existing digital preservation technologies or tools and our goal of establishing dynamic preservation.

A KM methodology is proposed and developed, and the utilizations of this methodology are:

- Drawing a whole guideline for development of KM in long term knowledge preservation projects;
- Identifying a hierarchical structure of KM;
- Identifying a sequential deployment process for KM;
- Designing data and process models for each level and each step.

One of our objectives, which are discussed in Chapter 1, is to propose a methodology, which is based on KM methodology, in order to carry out our research work. We propose a worksheet formula, in order to analyze and evaluate the existing KM methodologies. Although the KM methodologies are performed in different procedures and focus on different dimensions, we intend to use this worksheet formula to collect the information that we are interested in for our research work. We are going to separate the KM methodologies into different functional modules, and a functional module can be a process, a sub-process, or a simple activity. The functional modules are integrated and interacted in one KM methodology. By filling the worksheet formula (Table 2.1) that we have defined, we will have a comprehensive understanding of these KM methodologies in the viewpoint of functionalities. Because we use model-based development process to carry out our research work, the results of this functional module analysis of KM methodologies will help us to initialize the required functional modules of our own KM methodology, as well as to choose a KM methodology as our initial reference. The initial referential KM methodology will be combined with the extra functional modules. The worksheet formula we use to summarize the KM methodologies is shown in Table 2.1. The purpose of each column in the worksheet formula is presented, too.
<table>
<thead>
<tr>
<th>Number</th>
<th>Processes/Activities</th>
<th>How to/Phases</th>
<th>Outputs</th>
<th>Why/Objectives</th>
<th>Tool(s)</th>
<th>Deployment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>……</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Worksheet formula for analyzing KM methodologies

- **Number**: The number to identify each process in the methodology.
- **Processes/Activities**: The general processes or activities, which are proposed by the methodology.
- **How to/Phases**: The detailed description of one process/activity. All we consider this part as decomposition of one process/activity. In most of the KM methodologies, the processes and sub-processes are well illustrated.
- **Outputs**: The outputs of each phase. The output could be specific documents or signals or triggers for other activities or phases.
- **Why/Objectives**: The objective of one phase, and also the reason why we should add this phase in the methodology.
- **Tool(s)**: The supporting tool(s) or software that would will achieve the methodology.
- **Deployment**: The deployment of the KM phase or the supporting tool(s).

Of course not all the KM methodologies are described in a sequential manner, but we have utilized this formula to synthesize them, in order to identify the common functionalities as well as special features in KM (Annex 1). At the end of this synthesis, we are going to propose a methodology. Table 2.2 shows one example of one of the methodologies we have analyzed – CommonKADS methodology.
<table>
<thead>
<tr>
<th>Number</th>
<th>Processes /Activities</th>
<th>How to /Phases</th>
<th>Outputs</th>
<th>Why /Objectives</th>
<th>Tool(s)</th>
<th>Deployment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>Context level</td>
<td>Coping and feasibility study: analysis</td>
<td>OM: 5 worksheets</td>
<td>Identify problem/opportunity areas and potential solutions; Put them into a wider organizational perspective.</td>
<td>Organization Model</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coping and feasibility study: synthesis</td>
<td>OM: 5 worksheets</td>
<td>Decide about economic, technical and project feasibility; Select the most promising focus area and target solution.</td>
<td>Organization Model</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Impact and improvement study: analysis</td>
<td>TM: 2 worksheets; AM: 1 worksheet</td>
<td>study interrelationships between the task, agents involved, and use of knowledge for successful performance; what improvements may be achieved here.</td>
<td>Task Model, Agent Model</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Impact and improvement study: synthesis</td>
<td>TM: 2 worksheets; AM: 1 worksheet</td>
<td>Decide about organizational measures and task changes; Ensure organizational acceptance and integration of a knowledge system solution</td>
<td>Task Model, Agent Model</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>Summary: 1 worksheet</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Step 2</td>
<td>Concept level</td>
<td>Knowledge identification</td>
<td>Knowledge Model</td>
<td>survey the knowledge items; prepare them for specification</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Knowledge specification</td>
<td>Knowledge Model</td>
<td>complete specification of knowledge except for contents of domain models</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Knowledge refinement</td>
<td>Knowledge Model</td>
<td>Validate knowledge model; Fill contents of knowledge bases</td>
<td>structured walk-throughs; software tools for checking the syntax and find missing parts; paper-based simulation; prototype system</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CM: 2 worksheets</td>
<td>specifies knowledge/information transfer procedures; top-level control over task execution; additional communication tasks</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>Processes /Activities</td>
<td>How to /Phases</td>
<td>Outputs</td>
<td>Why /Objectives</td>
<td>Tool(s)</td>
<td>Deployment</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>---------</td>
<td>----------------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>Step 3</td>
<td>Artefact level</td>
<td>-</td>
<td>specification of a software architecture; design of the application within this architecture</td>
<td>Specify the architecture of implementation of the KM project</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2.2: Analysis result – CommonKADS methodology
From the analysis results of the CommonKADS methodology, we notice that there are three major functional modules, which is separated by different levels of viewpoints. In Context level and Concept level, there are decompositions and work processes, and each phase produces one or several worksheets of knowledge. On the contrary, the Artefact level of CommonKADS is not described in detail, and depends on the implementation methods and tools that the specific users of CommonKADS want to use. The tools that are utilized in each phase are predefined models (i.e. Organization Model, Task Model, etc.). The deployment of each phase is not specified in CommonKADS methodology, and some of the activities are performed without specific tools. Thus parts of the worksheet formula are blank. The detailed explanations of each KM methodology are shown in Annex1, and here we just discuss the results of our functional analysis of the KM methodologies. We illustrate the KM methodologies that we use in our analysis.

- **Accelerated Knowledge Management (AKM)** [Balafas et al. 03]: aims to develop KM methodology for long term perspectives. This methodology has been carried out at the Danwood Group in Lincoln and in collaboration with the Department of Computer Science, Loughborough University.

- **Distributed Knowledge Management (DKM)** [Cuel 03; Schwotzer et al. 04]: tries to manage knowledge in an autonomous way and introduces a concept of Knowledge Node (KN), which could represent a potential knowledge model for long term knowledge preservation.

- **MASK** [Ermine et al. 96; Barthelme et al. 98; Benmahamed et al. 05]: provides comprehensive approaches of knowledge capitalization: knowledge analysis and modeling. And the efficiency of the MASK methodology is proved by many KM projects.

- **MaKE** [Sharp et al. 03]: concerns the Information System (IS) development with KM approaches.

- **SAKE** [Ntioudis et al. 07-4, Ntioudis et al. 07-6]: describes the whole KM/KE approaches from analysis to deployment in the public administration environment. The approaches it describes and the connections between each step in KM/KE are helpful in our research work, which aims to develop an architecture and methodology for long term knowledge preservation.

- **SMARTVision** [Bubenstein-Montano et al. 01]: presents a more micro-view of a specific KM methodology, based on the existing KM methodologies. Different from
some existing KM methodologies, which are more strategic, SMARTVision provides a detailed description of each step of implementation of KM in designated organizations. The research on SMARTVision is helpful for implementing KM methodology in our research work, and contributes to discovering more threats and challenges in deploying a KM methodology into applications.

- **KM-Beat-It [Bures 05]:** is quite similar to the common KM methodologies. However, in the development of KM-Beat-It, some strengths and weaknesses of KM implementation are stated as a basis for this methodology. Thus KM-Beat-It really considers the issues of KM implementation, which are also our concerns in long term knowledge preservation projects.

- **Common Knowledge Acquisition and Design System (CommonKADS) [Orsvarn et al. 95]:** offers a structured approach to break down and structure knowledge engineering process. CommonKADS provides model-sets for creating requirements specifications for knowledge systems.

The results of the KM methodologies analysis lead to several sequential general functional aspects that KM should consider. They are:

- **Strategy alignment:** KM strategy must be predefined when starting a KM project in designated organizations. This strategy will lead to the following steps or functions in one KM project.

- **Knowledge identification:** to identify existing data, information and knowledge from reliable sources within or across organizations.

- **Knowledge acquisition:** based on the knowledge identification and knowledge source identification, the procedures/functions/technologies for knowledge acquisition will be determined and performed.

- **Knowledge modeling:** a knowledge model is an interpretable model of knowledge or standard specifications about a kind of knowledge.

- **Knowledge adaptation:** after knowledge has been extracted and encapsulated thanks to the knowledge models, knowledge should be stored in repositories, as the security of the preserved knowledge must be ensured.

- **Knowledge transfer:** the process and tools of knowledge transfer should be identified. The tools for transferring knowledge could be in the form of a static network approved by the designated organization, or web services, etc.
• Knowledge evaluation: the dynamic analysis and identification of knowledge.
• Knowledge Revise: the dynamic maintenance of knowledge.

These aspects appear in nearly all the KM methodologies. Through our analysis results, these aspects together constitute the basics of the KM architecture. The identification of these fundamental aspects constitutes the initial work for our research. If certain tools provided the functionalities that achieve these aspects, it would be quite possible to use the very tools to support a KM project. And moreover, if these tools, which are with KM functionalities, collaborate, we would establish an integrated KM system.

### 2.2.3. Knowledge Management Tools

We try to find out if the fundamental aspects of KM are supported by existing tools, thus we have performed on the analysis in existing commercial software tools, which concern data/information/knowledge management. From the research and analysis of [Banerjee 04], the KM tools are divided into 4 business classes and 16 functional sub-classes:

<table>
<thead>
<tr>
<th>Technology</th>
<th>Tool Type / Functional Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge Generation</td>
<td>Meeting Support Tools</td>
</tr>
<tr>
<td></td>
<td>Structuring</td>
</tr>
<tr>
<td></td>
<td>Visualizing</td>
</tr>
<tr>
<td></td>
<td>Polling</td>
</tr>
<tr>
<td></td>
<td>Group Decision Support Software</td>
</tr>
<tr>
<td></td>
<td>Data-mining</td>
</tr>
<tr>
<td></td>
<td>Online Analytical Processing (OLAP)</td>
</tr>
<tr>
<td>Knowledge Codification</td>
<td>Knowledge Repositories</td>
</tr>
<tr>
<td></td>
<td>Document Management Systems</td>
</tr>
<tr>
<td></td>
<td>Text-mining</td>
</tr>
<tr>
<td></td>
<td>Taxonomy Generators</td>
</tr>
<tr>
<td>Knowledge Retrieval</td>
<td>Retrieval Systems</td>
</tr>
<tr>
<td></td>
<td>Search Machines</td>
</tr>
<tr>
<td></td>
<td>Navigators</td>
</tr>
<tr>
<td>Knowledge Transfer</td>
<td>Online Collaboration</td>
</tr>
<tr>
<td></td>
<td>Online Coordination</td>
</tr>
<tr>
<td></td>
<td>Training Tools</td>
</tr>
</tbody>
</table>

Table 2.3: Classification of KM tools [Banerjee 04]
In order to prove this taxonomy, we have carried out a functional analysis of 78 KM software tools, which are developed by 33 different software companies. The result of the analysis shows that the taxonomy of KM tools is fair (Figure 2.1, Annex 2).

![Figure 2.1: Distribution of KM tools by functional taxonomy of Banerjee](image)

This taxonomy is fair and it seems that, on the market, the distribution of each kind of commercial tools is similar. However, there are no KM tools that cover all or at least most of these features (Figure 2.2). So if we try to establish an integrated KM system, we need to get benefits of all sorts of features from different tools.

![Figure 2.2: Coverage of features by KM tools](image)
2.2.4. Knowledge Management System

Knowledge management systems (KMS) have been defined as “an emerging line of systems which target professional and managerial activities by focusing on creating, gathering, organizing and disseminating an organization’s ‘knowledge’ as opposed to ‘information’ or ‘data’ [Alavi et al. 99]. We consider that such a system is an enhanced digital information preservation system, which integrates technologies (e.g. artificial intelligent technologies [Becerra-Fernandez 00], etc.) to support modern knowledge management.

The KMS could be utilized in different ways as in different organizations. However, through the research and observation of Becerra-Fernandez [Becerra-Fernandez 99], KMS fall into three categories:

- **Educational KMS**: this type of KMS is usually an educational tool for training in organizations. However, the educational training is also one way of eliciting and inheriting tacit knowledge.

- **Problem-solving KMS**: this type of KMS is implemented with intellectual technologies, and aims to capture knowledge for reusing and solving problems. The problems concerning knowledge could be new problems, which need solutions derived from existing knowledge; and the problems could also be old problems, which need recurring to the existing solutions. In any condition, problem-solving KMS’ perspective is problem-oriented.

- **Knowledge repositories**: this type of KMS is the widest used in organizations. Although there are sub-categories under knowledge repositories, we just consider it as a preservation system, which keeps knowledge in certain designated forms.

We introduce the “system” concept, in order to establish a platform for the perspective of long term digital preservation. In other words, we are going to develop a digital preservation system, which implemented as repository as well as a problem-solving KMS. A digital preservation system refers to a system in which objects are stored for preservation archiving. Thus, we have to analyze the existing functionalities and technologies of digital preservation system/KMS, and to integrate the functionalities and technologies as one system in the perspective of long term knowledge preservation. The detailed architecture of our research work will be discussed in Chapter 4.
2.3. Proposal of KM Approach

CommonKADS stands for Common Knowledge Acquisition and Design System which is a present version of KADS. CommonKADS is one of the KM methodologies that we have analyzed in the section “2.2.2. Knowledge Management Methodologies”. It enables a top-down approach and provides model sets at each level. The reason why we have chosen CommonKADS as our KE approach for long term knowledge preservation is that we intend to get benefits of its abundant models. We are going to develop an approach and an architecture in the perspective of model reusability in the long term. The top-down structure and model-sets composition of CommonKADS are fit for our needs. Therefore, we choose CommonKADS from other KM methodologies, because we will benefit from its model sets and KADS templates to develop a model-based architecture.

The method has been developed since 1984 through two major CEC ESPRIT (Commission of the European Communities, European Strategic Programme for Research and development in Information Technologies) funded research projects. The methodology aims to support structured knowledge engineering. It indicates the opportunities and bottlenecks in the organizations, distributes and applies their knowledge resources, and so gives tools for corporate knowledge management. It also provides the methods for performing a detailed analysis of knowledge-intensive tasks and processes. CommonKADS supports the development of knowledge systems that support selected parts of the business process [Schreiber et al. 99].
The three levels of CommonKADS methodology are:

- **Context Level** analyzes the organizational environment and the corresponding critical success factors for a knowledge system. In other words, the model and functionalities in this level interact with the knowledge source (i.e. data, information in designated organization).

- **Concept Level** yields the conceptual description of problem-solving functions and data that were handled and delivered by a knowledge system. The models and functionalities in this level manipulate the knowledge, which is acquired from knowledge source.

- **Artifact level** integrates the above levels together in the design model in order to construct the requirements specification for the knowledge system. In original CommonKADS methodology, this level is not specifically defined.

CommonKADS is a complete methodological framework for the development of a Knowledge Based System (KBS) [Schreiber et al. 99]. It supports most aspects of a KBS development project, such as:

- Project management
- Organizational analysis (including problem/opportunity identification)
- Knowledge acquisition (including initial project scoping)
- Knowledge analysis and modeling
- Capture of user requirements
- Analysis of system integration issues
- Knowledge system design

We have proposed the KM methodology [Teng et al. 10], which is an extension of CommonKADS. We have introduced the Product-Process-Organization (PPO) concept in our proposed KM methodology, in order to acquire production related knowledge. In general, comparing to the three levels of the original CommonKADS methodology, the approaches of the methodology are divided into 4 sequential modules (levels):

- **Context Level**: this is the knowledge identification phase. In this phase, we must capture a corporate organization structure and strategy of digital preservation in enterprise. Then simultaneously, we identify the business processes and product
knowledge. Here the “product” not only refers to the product in the manufacturing industry, but any kind of product, which uses an information system and digital descriptions. We have introduced PPO concept in our KM methodology, thus in Context Level, we have to modify the original models into Organization Model, Task Model and Product Model, in order to construct the models of knowledge acquisition dedicated to product and production related information. Therefore, we added a Product Model. The original Agent Model is canceled in our proposal because we don’t require too much information of the operators in production workshop.

- **Concept Level:** this is the knowledge integration phase. In this phase, knowledge is manipulated according to the knowledge models and templates we have defined. The models and templates are defined in the perspective of long term preservation. The knowledge captured in Context Level will have to be formalized in order to archive. As our perspective of this research work is long term preservation, we have to consider the digital preservation issue even we are developing the KM methodology. In Concept level, we redesign the Knowledge Model to fit for digital preservation. We also propose a Transformation Model, which is used for constructing and transferring knowledge between different models. The original Communication Model is canceled in our proposal because we have already canceled the Agent Model and Communication Model describes the communication plan between agents.

- **Design Level and Implementation Level:** In Design Level, we have to synthesize the results of the previous phases, and we design the architecture of long term knowledge preservation from the business point of view. Inside of this architecture, we have to identify and design functionalities and sub-functionalities for each business process. In Implementation Level, we implement the previous designs. The reason we add design and implementation model is that the original CommonKADS methodology provides little instruction on its Artefact Level, but in our research work, we need to propose an architecture and functional designs for long term preservation. Therefore, we replaced the original Artefact Level by Design Level and Implementation Level, in order to provide more detailed development instruction for deploying the KM methodology.

The general view of our extended CommonKADS methodology is shown in Figure 2.4. We have marked the models that are modified or added.
From Figure 2.4, we notice that, compared to the original CommonKADS, some models are replaced by new models. Even for the models that are not replaced, the detailed identification and design of models will be different from CommonKADS. The adaptation and extension of CommonKADS can be considered as:

- Adaptation of the general structure of CommonKADS, or in other words, the way that CommonKADS manages knowledge related projects.
- Adaptation of certain models (e.g. Organization Model), and extension of some models (e.g. Knowledge Model). The extension means that the model may be similar to the original model, but some features are changed by our needs.
- Introduction of new levels and models (e.g. Architecture Model).

In our research work, we work on the long term knowledge preservation with production information. This is also the reason why we have introduced the PPO (Product-Process-Organization) model in the Context Level of our proposal, although the “product” can be extended to other products rather than manufacturing products. However, in this thesis, we use production information systems for our research work. In this chapter, we introduce the utilization of each level and model of our proposed KM methodology. The detailed design of
architecture and functionalities will be introduced in Chapter 4. And the implementation of our design will be introduced in Chapter 5.

2.3.1. Context Level Design

In production information systems, the management and maintenance of product structure is one of the most important functions in the whole system [Saaksvuori 08]. For example, modern PLM (Product Lifecycle Management) systems handle several product structures for the same product from different viewpoints (e.g. the product structure is different when considered from the engineering point of view than a manufacturing point of view). Thus, we propose to formalize the product related knowledge by mapping from data in production information systems to the PPO model. The reason why we have had this proposal is that data are not always managed by using PPO model in information systems in enterprise. In our proposed methodology, the modeling process is supported by the following models/documents:

- Organization Model (OM) is the scope and feasibility study which describes and analyzes a broader organizational environment: 5 worksheets. The Organization Model represents the organization environment for knowledge management in an enterprise;
- Task Model (TM) focuses on tasks (sub-processes, breakdown of business processes) of PLM and identifies the information systems that operate the tasks: 2 worksheets. The Task Model represents the real tasks performed in enterprise, not the KM tasks;
- Product Model (PM) collects all product related information: 1 document package, whose format depends on the product engineering data representation and exchanging implementation methods.

In a long term knowledge preservation context, some original CommonKADS models will not be promptly applied. And we must redesign the context models.

2.3.1.1. Organization Model

Organization Model supports the analysis of the major features of an organization, in order to discover problems and opportunities for the knowledge system, establish their feasibility, and assess the impacts on the organization of intended knowledge actions. In a long term
knowledge preservation context, Organization Level is one significant aspect in PPO model design; therefore, we keep the Organization Model of CommonKADS in the model set.

The worksheets OM-1 to OM-5 are used for interviewing knowledge decision makers, who are responsible for a KM approach, in organizations, or analyze the information system organizational structure in the enterprise. Then, the outputs from the model are the list of the knowledge intensive processes and product knowledge assets which are related to each process. Finally, the feasibility of the knowledge management project was analyzed to see if the project was feasible in terms of business, technique, project and solution. It serves as a decision support for the study of business, technical and project feasibilities, in order to select the most promising focused area and targeted solution [Sureephong 09]. The five worksheets are show in Figure 2.5 (OM-1 to OM-5).

![Figure 2.5: Organization Model worksheets](image)

In general, the OM worksheets focus on the products, respecting the CommonKADS methodology.

- OM-1 analyzes the KM project problem (e.g. long term digital preservation) and organizational context (i.e. enterprise strategy, goals, missions, and important external factors, etc.), and lists suitable solutions and technologies that could be adapted.
- OM-2 represents a single problem solution of OM-1, and contains information regarding the organizational structure, business process, product and knowledge.
OM-3 identifies the business process, and breaks down the process, which is concerned in OM-2, into tasks (sub-processes). Simultaneously, OM-3 identifies the corresponding products and information systems of the tasks.

OM-4 identifies each product, introduced in OM-2 and OM-3. OM-4 specifies the corresponding information system as the source of product knowledge.

OM-5 is a decision-making support document, which summarizes the worksheets from OM-1 to OM-4, and focuses on business, technical and project feasibilities. Then it proposes actions, risks and constraints of the KM project.

### 2.3.1.2. Task Model

For enterprises and industries, an archive must capture all the data required to completely define the product, and in some instances, processes [Lubell et al. 09]. Task Models are the relevant subparts of a business process. The Task Model analyzes the global task layout, its input and outputs, preconditions and performance criteria, as well as needed resources and competences skills. We reform the Task Model by reducing the knowledge item description and adding the modeling of executors of processes. The Task Model can be considered as a process model or sub-process model, corresponding to the PPO design.

The Task Model is a refinement of knowledge intensive tasks identified in the Organization Level. To investigate a task, three viewpoints are concerned in this model. The functional view divides a task into subtasks: input and output. The static information structure view is a description of the information content and structure of objects that are handled in the task. The control view (or dynamic view) provides understanding about triggering events, decision-making points, and other knowledge about the time aspects. The two worksheets are shown in Figure 2.6 (TM-1, TM-2).

- TM-1 aims at refining the task within the target process. The three views (i.e. functional view, static information structure view and control view) of tasks are addressed by this worksheet.

- TM-2 is a specification of the information system or sub-functions of an information system, where the target task performs. This worksheet, which concerns information system, is quite different from the original CommonKADS methodology, because we propose a methodology that is dedicated to production related digital preservation.
2.3.1.3. Product Model

The PM-1 document is a specification of the product knowledge employed for a task, and possible bottlenecks and areas for improvement. In information systems of production, the Product Model is obvious and easy-to-captured. In fact in most information systems, there are already specific data structures for product, especially in PDM (product data management) or PLM systems. The composition of Product Models will cover the following aspects: component, function, behavior, structure, interface, specification and metadata, but will not be limited to these aspects. This worksheet structure depends on the product data structure in information systems.

Generally, the Context Level models collect information from information systems, which are the knowledge source and the results of the Context Level models (especially Organization Model) determine the feasibility of the long term knowledge preservation research. We have identified the correlations between the models in Context Level (Figure 2.8). Task Model is the decomposition of the OM-3, and Product Model is the decomposition of OM-4. Some parts of our proposed structure are from the original CommonKADS methodology (e.g. correlation between OM-1 and OM-5), but the detailed worksheet of each model and more other correlations between OM, TM and PM are proposed according to PPO design concept.
2.3.2. Concept Level Design

The worksheets in the Context Level act as checklist and information archive, and they should be used with flexibility. In the Concept Level we have to identify models, which are used for manipulating the information archive.

In order to establish the digital preservation platform, for knowledge archiving, we have adapted the OAIS reference model. Thus, in the digital preservation platform, knowledge
appears as the form of Information Package (IP), which is the knowledge model provided in OAIS. Therefore, the extracted knowledge from information systems will finally be packaged in the Information Package form for long term knowledge preservation. In other words, the Knowledge Model will be constructed according to the Information Package structure.

### 2.3.2.1. Knowledge Model

The knowledge model is the formalization of the knowledge, which is acquired from Context Level. In other words, the knowledge model is a comprehensive knowledge object. Our proposed KM methodology targets long term digital preservation, thus the knowledge model is structured according to the digital preservation aspects. We have adapted the OAIS reference model to establish a digital preservation platform. Consequently, we adapt the knowledge object structure in OAIS, and the Knowledge Model in our methodology is an extension of Information Package (IP) model from OAIS. IP is a package containing data objects as well as comprehensive metadata to describe the data objects (Figure 2.9).

![Figure 2.9: Information Package concepts and relationships [CCSDS 650.0-B-1; ISO 14721:2003]](image)

The detailed explanations of the concepts of IP are illustrated as follows:

- Content Data Object: Physical Objects or Digital Objects.
• Representation Information: Information that makes the Content Data Object understandable to the Designated Community.

• Structure Information: Information to describe structure or data format.

• Semantic Information: Semantic description, varied and complex.

• Reference to the other Representation Information: When Representation Information itself is an Information Object, it needs other Representation Information to explain its own Digital Objects.

• Preservation Description Information (PDI): Information to preserve the Content Information, to ensure it is clearly identified, and to understand the environment in which the Content Information was created.

• Provenance Information: Source and history of Content Information.

• Context Information: Relationship of Content Information and other information outside the package.

• Reference Information: Identifiers or systems of identifiers to identify the Content Information uniquely.

• Fixity Information: Wrapper of protective shield to protect the Content Information from undocumented alteration.

• Package Information: Information which, either actually or logically, binds, identifies and relates the Content Information and PDI.

• Descriptive Information: Information to discover which package has the Content Information of interest.

The product data will be preserved as data objects, while the IP itself is task-oriented. In Descriptive Information of the IP, the usage of the IP (i.e. task objective of IP) is described based on the analysis in Task Model. Thus the end users of the preserved IP will locate the corresponding knowledge by their working requirements. For example, if one end user tries to find information about core design of a power transformer in the critical design phase, his/her task objective is “core design”. And by searching Descriptive Information with this task objective, all the corresponding IPs would be located. Figure 2.7 is just a logical view of the information package, and Figure 2.10 shows the class diagram of the knowledge model: Information Package.
2.3.2.2. Transformation Model

The Transformation Model in our KM methodology focuses on knowledge model-transfer, in other words, the knowledge mapping from one sort of model to another. The knowledge creation process in Context Level is in a top-down strategy. Thus, either Task Model or Product Model is identified in Organization Model. And in this manner, it is not too difficult to establish links to merge the Context models (Figure 2.6). Then the outcome of the Context Level models and worksheets will provide not only a clear idea on the target knowledge for long term preservation, but also a comprehensive knowledge of the organization, business process and product engineering. In other words, the Transformation Model includes the mapping from context model sets to formalized knowledge model. However, the knowledge transfer process is not always an integration process. For instance, in knowledge reuse process, specific information system may require just product knowledge, which needs to be separated from the content in Knowledge Model.

In digital preservation platforms, the worksheets are easy to convert from or to XML (Extensible Markup Language) format. We use XML as one example to explain how to identify the communication mapping.

Figure 2.10: Knowledge Model: Information Package
The Product Model document may also be represented in XML. By using XML, we integrate the product knowledge with organization and process knowledge, in order to capture comprehensive metadata for long term preservation. According to the KM approach and worksheets structure, we have developed our XML Schema for generating XML files from worksheets. We provide these XML files as the knowledge source of construction of knowledge model. Figure 2.11 shows the mapping from Context models to knowledge model.

The knowledge model mapping (conversion) in Figure 2.11 is from the left side to the right side. According to this mapping, not all the worksheets are converted into XML. OM-1 and OM-5 are decision-making documents in our KM project, while other six documents are the source knowledge for integration. Thus, the source knowledge documents are converted into XML files and are put in XML database, waiting for more operations. Based on the IP structure, we query corresponding information and compose the information as the components in IP. All the components are generated into two XML files (i.e. Preservation Description Information and Representation Information), which is the metadata of the product data object. And at the same time, the product data object is also saved. The linking information of product data and metadata is described in Packing Information. Simultaneously, the Descriptive Information is generated, too.

Figure 2.11 just shows one example of knowledge model mapping. In fact, the knowledge model mapping (conversion) rules includes certain terms and scenarios that will decide when and how the knowledge model should convert from one to another.
Figure 2.11: Transformation Model: mapping from Context models to Knowledge Model
2.3.3. Discussion of Design and Implementation Levels

It is argued that, the Artefact level (i.e., the design models) was lightly defined in the CommonKADS methodology. Hence, in our research work, software engineering concept is introduced to extend the CommonKADS and two levels (Design Level and Implementation Level) are proposed to replace the Artefact level.

In Design Level, Architecture Model should be developed. Not only the structure of long term knowledge preservation is established, but also specifications of required functionalities and services are produced in Architecture Model. Therefore, the Architecture Model and Function Model include not only the KM activities, but also the knowledge preservation activities. We will introduce the architecture and the required functional designs after the discussion of digital preservation platform (Chapter 3), where knowledge preservation is performed. The Design Level (i.e. Architecture Model and Functional Model) will be introduced in Chapter 4. We have noticed that in our research work, the extended CommonKADS methodology is the direction and blueprint for not just KM approach, but for the whole long term knowledge preservation approach. Nevertheless, during the discussion from Chapter 2 to Chapter 3, we will see some general design of Function Model, and in Chapter 4, they are discussed in details.

The last level is Implementation Level is the deployment of KM architecture and models. The level translates the architecture and models into software programs. The Deployment Model depends on the models we have discussed above as well as the software platform. The research on digital preservation platform is one other major work of our research. Although the CommonKADS has this hierarchic structure, the Architecture Model or service model design depends on, to certain extent, the selection of digital preservation platform. The Implementation Level is the lowest in the hierarchic structure, but the Design Level needs the feedback from this level for architecture and service design. In fact the implementations of KM approach and digital preservation approach ought to be integrated, because they will not realize long term knowledge preservation by either one alone. Thus, we will introduce the implementation approaches in detail in Chapter 5.
2.3.4. Model Sets in KM Approach

We have already described the model sets of our CommonKADS methodology. These models are the basics of our research work. These predefined models will be used in each step of the long term preservation processes. In Section 2.2.2, we have already discussed the major steps of KM approach. Accordingly, in our research work, we have identified the functions in the processes of knowledge acquisition and knowledge reuse in enterprise. Table 2.4 shows the general description of the Function Model of the KM approach in our research work. In Table 2.4, we notice that the Context Level models are required to associate with each function or sub-function. In other words, the functional modules or KM approach depend on the Context Level models.

<table>
<thead>
<tr>
<th>General Info.</th>
<th>Function</th>
<th>Sub-function</th>
<th>Associated Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>KM Approach</td>
<td>KM Planning</td>
<td>Develop KM Strategy</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Capture Critical Change</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Design Knowledge Model</td>
<td>Organization Model</td>
</tr>
<tr>
<td></td>
<td>Knowledge</td>
<td>Capture Organizational Info.</td>
<td>Organization Model</td>
</tr>
<tr>
<td></td>
<td>Acquisition</td>
<td>Identify Business Process</td>
<td>Organization Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identify Product Info.</td>
<td>Organization Model</td>
</tr>
<tr>
<td></td>
<td>Knowledge</td>
<td>Handle Query</td>
<td>Organization Model</td>
</tr>
<tr>
<td></td>
<td>Evaluation</td>
<td>Determine Knowledge</td>
<td>Task Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Obsolescence</td>
<td>Product Model</td>
</tr>
<tr>
<td></td>
<td>Reuse</td>
<td>Create Query</td>
<td>Organization Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distribute Knowledge</td>
<td>Task Model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Product Model</td>
</tr>
</tbody>
</table>

Table 2.4: Function Model of KM approach

The functions related to KM approach in enterprise:

- KM Planning: the KM planning is the decision making part of KM. According to different environments (e.g. different information systems used in different departments, etc.), different cultural issues would be captured, and consequently the KM objectives and KM methods would be determined (sub-function: Develop KM strategy). This function also designs the structure of the extracted knowledge from information systems and domain experts (sub-function: Design Knowledge Model)
and identifies the changes (sub-function: Capture Critical Change) that would trigger the knowledge changes in digital preservation platform.

- Knowledge Acquisition: this function extracts data and knowledge from information systems or from domain experts, according to the knowledge model, which has been designed in KM planning function. And the packaged knowledge is transferred into the digital preservation platform.

- Knowledge Evaluation: the Enterprise Layer also receives queries when knowledge repository performs the knowledge obsolescence check. Knowledge evaluation function handles this kind of queries and makes decision whether the preserved knowledge should be kept or not, according to the query results. This function aims to reduce the redundancy of the preservation.

- Knowledge Reuse: this function plays as a contrary role as the Knowledge Acquisition. Firstly, this function sends knowledge reuse request to the knowledge repository. Then it gets the responses and results from the knowledge repository, whether appropriate knowledge is retrieved or not. If the appropriate knowledge is sent back, this function sends the knowledge back to information system or domain expert for reuse.

2.4. Conclusion

We have presented, in this chapter, the KM methodology for our research work. We develop our KM methodology by extending the CommonKADS methodology. And our proposed methodology has a multiple levels structure and top-down approach, with several model sets. We have introduced each part of our proposed KM methodology level by level in this chapter. The methodology describes from the KM approaches in knowledge sources context, to knowledge modeling and communication activities among information systems. Although we have identified the functionalities of KM approach, the detailed functional design would be discussed in Chapter 5 after the presentation of the global architecture of our long term knowledge preservation research work.