Chapter-6

COMBINING MULTIPLE KNOWLEDGE SOURCES

Knowledge is very distinct from data and information and provides a higher level of meaning about that data and information. The ability to act is an integral part of being knowledgeable. **Data** are a collection of Facts, Measurements, and Statistics. **Information** is organized or processed data that are Timely and Accurate. **Knowledge** is information that is Contextual, Relevant and Actionable. Having knowledge implies that it can be exercised to solve a problem, whereas having information does not.

Knowledge has the following characteristics that differentiates it from an organization’s other assets

- Extraordinary leverage and increasing returns. Knowledge is not subject to diminishing returns. When it is used, it is not consumed. Its consumers can add to it, thus increasing its value.

- Fragmentation, leakage, and the need to refresh. As knowledge grows, it branches and fragments. Knowledge is dynamic; it is information in action. Thus, an organization must continually refresh its knowledge base to maintain it as a source of competitive advantage.

- Uncertain value. It is difficult to estimate the impact of an investment in knowledge. There are too many intangible aspects.

- Uncertain value of sharing. Similarly, it is difficult to estimate the value of sharing the knowledge, or even who will benefit most.

- Rooted in time.
• Combining multiple knowledge sources certainly enhances Intellectual capital or intellectual assets.

• Tacit knowledge is usually in the domain of subjective, cognitive, and experiential learning; it is highly personal and difficult to formalize. It is also referred to as embedded knowledge since it is usually either localized within the brain of an individual or embedded in the group interactions within a department or business unit.

• Tacit knowledge is the cumulative store
  1. of the corporate experiences
  2. Mental maps
  3. Insights
  4. Acumen
  5. Expertise
  6. Know-how
  7. Trade secrets
  8. Skill sets
  9. Learning of an organization
  10. The organizational culture

**Tacit knowledge is generally slow and costly to transfer and can be plagued by ambiguity.**

The goal of knowledge management is for an organization to be aware of individual and collective knowledge so that it may make the most effective use of the knowledge it has. Firms recognize the need to integrate both explicit and tacit knowledge into formal information systems - Knowledge Management System (KMS)

### 6.1 Knowledge management (KM):

**Knowledge Management** is a process that helps organizations identify, select, organize, disseminate, and transfer important information and expertise that are part of the organization’s memory. Structuring of knowledge enables effective and efficient problem solving, dynamic
learning, strategic planning and decision making tasks easier. Knowledge management initiatives not only focus on identifying knowledge but also on how it can be shared in a formal manner by leveraging its value through reuse. Knowledge management can promote organizational learning thereby helping the organization to solve problems systematically.

A functioning knowledge management system follows six steps in cycle dynamically refining information over time.

1. Create knowledge.
2. Capture knowledge.
3. Refine knowledge.
4. Store knowledge.
5. Manage knowledge.
6. Disseminate knowledge

As knowledge is disseminated, individuals develop, create, and identify new knowledge or update old knowledge, which they replenish into the system.

**Fig 6.1 Example of – Knowledge Management Systems (WSD)**

- Combined Multiple Knowledge sources management initiatives have one of three aims:
  1. To make knowledge visible mainly through Maps, yellow pages, hypertext.
  2. To develop a knowledge-intensive culture,
3. To build a knowledge infrastructure

There are several activities or processes that surround the management of combined knowledge.

- Knowledge Creation
- Knowledge Sharing
- Knowledge Seeking

### 6.2 Knowledge creation or knowledge acquisition

The generation of new insights, ideas, or routines.

- **Socialization mode** refers to the conversion of tacit knowledge to new tacit knowledge through social interactions and shared experience.

- **Combination mode** refers to the creation of new explicit knowledge by merging, categorizing, reclassifying, and synthesizing existing explicit knowledge.

- **Externalization** refers to converting tacit knowledge to new explicit knowledge.

- **Internalization** refers to the creation of new tacit knowledge from explicit knowledge.

- **Knowledge sharing** is the exchange of ideas, insights, solutions, experiences to another individuals via knowledge transfer computer systems or other non-IS methods.

- **Knowledge seeking** is the search for and use of internal organizational knowledge.

- There are two fundamental approaches to knowledge management: a process and a practice approach. Since the two are not mutually exclusive a knowledge management initiative will probably involve both approaches.

- The process approach attempts to codify organizational knowledge through formalized controls, processes, and technologies frequently through the use of information
technologies to enhance the quality and speed of knowledge creation and distribution. These technologies include:

a. Intranets
b. data warehousing
c. knowledge repositories
d. decision support tools
e. groupware

- The process approach is favored by firms that sell relatively standardized products since the knowledge in these firms is fairly explicit because of the nature of the products & services.

- The practice approach to knowledge management assumes that organizational knowledge is tacit in nature and formal controls, processes, and technologies are not suitable for transmitting this type of understanding.

- Rather than building formal systems to manage knowledge, this approach builds social environments or communities to facilitate the sharing of tacit understanding.

- The practice approach is typically adopted by companies that provide highly customized solutions to unique problems. The valuable knowledge for these firms is tacit in nature, which is difficult to express, capture, and manage.

- Enhancement can be made using WSD feature in Most knowledge management software packages include one or more of the following tools:
  a. collaborative computing tools
  b. knowledge servers
  c. enterprise knowledge portals
  d. electronic document management systems
Technology tools that support knowledge management are called knowledge harvesting tools. Technologies enable advanced functionality by combining multiple knowledge management systems and form the base for future innovations.

- **Artificial Intelligence (AI methods: expert systems, neural networks, fuzzy logic, genetic algorithms, etc.)**
  
  a. Assist in identifying expertise
  b. Elicit knowledge automatically and semi-automatically
  c. Provide interfacing through natural language processors
  d. Enable intelligent searches through intelligent agents.

- **Intelligent agents are software systems that learn how users work and provide assistance in their daily tasks.**

- **WSD based Knowledge Discovery in Databases (KDD) is a process used to search for and extract useful information from volumes of documents and data. It includes tasks such as:**
  
  a. knowledge extraction
  b. data archaeology
  c. data exploration
  d. data pattern processing
  e. data dredging
  f. information harvesting

- **WSD benefits Consulting Firms provide assistance**
  
  a. in establishing knowledge management systems
  b. measuring their effectiveness
c. Support for vertical market software
d. Application service providers (ASPs) with Disambiguation have evolved as a form of KMS outsourcing on the Web. Offering a complete knowledge management solution, including a KM suite and the consulting to set it up.

6.3 Knowledge Management

Integration frame work helps to build enterprise solutions.

- Combined Knowledge management from multiple sources systems are enterprise-wide and must be integrated with other information systems in an organization. The benefits include

  a. Decision Support Systems (DSS)
  b. Artificial Intelligence
  c. Customer Relationship Management Systems (CRM)
  d. Supply Chain Management Systems (SCM)
  e. Corporate Intranets
  f. Extranets

Knowledge Management – Integration can be represented as given below.
Fig 6.2 Multi perspective Knowledge integration framework
6.4 Beneficiaries of Knowledge Integration (using WSD cognitive approach) are

- Chief knowledge officer’s (CKO) role are to maximize the firm’s knowledge assets, design and implement knowledge management strategies, effectively exchange knowledge assets internally and externally, and promote system use.

- Chief executive officer’s (CEO) is responsible for championing the KM effort.

- Chief financial officer (CFO) must ensure that the financial resources are available.

- Chief operating officer (COO) must ensure that people begin to embed knowledge management practices into their daily work processes

- Chief information officer (CIO) is responsible for the IT vision of the organization and the IT architecture, including databases, application software, etc.

- KMS developers are the individuals who actually develop the system

- KMS staff catalogue and manage the knowledge, train users

6.5. Benefitting Managerial Issues by using integrated WSD approach of Knowledge Management are:

- **Organizational culture change**: This issue is how can one change organizational culture so that people are willing both to contribute knowledge to and use knowledge from a KMS? There must be strong executive leadership, clearly expressed goals, user involvement in the system, and deployment of an easy-to-use system that provides real value to employees. A viable reward structure for contributing and using knowledge can also be developed.
• **How to store tacit knowledge**: This is extremely difficult. Most KMSs (based on the network storage model) store explicit knowledge about the tacit knowledge that people possess. When the knowledgeable people leave an organization, they take their knowledge with them. Since knowledge requires active use by the recipient, it is important for the person generating knowledge to articulate it in a way that another, appropriately educated person can understand it.

• **How to measure the tangible and intangible benefits of KMS**: There are a number of ways to measure the value of intellectual assets and of providing them to the organization.

• **Helps in Determining** the roles of the various personnel in a KM effort.

Measuring success or Failures can be easily done by implementing this approach of WSD based knowledge management strategy. This valuation can be based upon an asset-based approach or one that links knowledge to its applications and business benefits.

  a. WSD based Asset-based approach starts with the identification of intellectual assets and then focuses management’s attention on increasing their value.
  b. WSD based variants of a balanced scorecard, where financial measures are balanced against customer, process, and innovation measures.
  c. Removing ambiguities in Financial Metrics (tangible benefits).
  d. Providing safety cover for Non-Financial Metrics (intangible benefits).

• **The lasting importance of Integrated knowledge management.** Knowledge management is extremely important. If it is correctly done, it can have massive impact by leveraging know-how throughout the organization. If it is not done, or is not correctly done, the company will not be able to effectively compete against another major player in the industry that does KM correctly.
Implementation of Integrated context, content and common knowledge based knowledge framework in the face of quickly changing technology. This is an important issue to address regarding the development of many IT systems. Technology has to be carefully examined, and experiments done, to determine what makes sense. By starting now, an organization can get past the managerial and behavioral issues, which have greater impact on the eventual success (or not) of a KMS. As better and cheaper technology is developed, the KMS can be migrated over to it, just as legacy systems have migrated to the PC.

6.6 Knowledge-based Content Navigation in e-Learning Applications (WSD approach)

In this context of thesis, an integrated framework for organizing and navigating online learning material with respect to the semantic context of documents is described. Here the use of a fuzzy clustering algorithm and Topic Maps to discover and represent knowledge considerably helps to resolve the ambiguity issues respectively.

WSD benefits to E-learning applications: Although e-Learning applications lack face-to-face interaction between teachers and students, they present the major advantage of enabling people to access learning facilities regardless of their location and at the time that is most convenient to them. Large networked repositories of learning material may be accessed by students, but WSD tools are needed to narrow down the available resources to a particular individual based on the learning context, i.e., take into account learning objectives, pedagogical approaches, user profile, etc. Hence, tools are required to determine which documents are the most relevant for a given student, who wants to learn a particular subject.

The Document relevance should be computed based on two components. Firstly, the set of concepts associated with each document should be identified. Thus, there needs to be a way to classify and organize materials in terms of knowledge domains. Secondly, the learning context should be considered, since the learning goals and pedagogical models may impact on the way materials are structured and consequently, on the definition of relevant links. Cognitive
integrated approach of WSD is to use fuzzy clustering to identify relationships between learning materials and to dynamically organize them into knowledge domains. Thus Knowledge integration of this standard helps in adaptive hypermedia systems to build customized learning sessions.

### 6.7 Dynamic Knowledge Discovery and Representation

Knowledge representation is becoming ever more important in applications that deal with large amounts data. The interpretation and sharing of data between systems in meaningful ways using automated tools requires knowledge about the data itself. It is therefore necessary to have semantic descriptions of the information so that knowledge can be generated and exchanged. Metadata (i.e. data about data) serves that purpose since it can provide rich descriptions about resources.

In e-Learning applications metadata is also fundamental for describing online learning materials and among other information, it can capture the subjects or knowledge domains associated with each document. This information is needed for customizing learning sessions, but it is not sufficient. For locating relevant content, the system needs to know how the subjects are interrelated in an abstract knowledge space and follow the associations to suggest links to the student. There are several approaches to build the knowledge space one of which is to manually create an ontology of the problem domain. But this approach as some limitations namely its maintenance effort and subjectivity. process of dynamic knowledge discovery and representation are hence handled using **fuzzy clustering**.

**Clustering:** Clustering algorithms aim at grouping data elements according to some similarity measure so that related elements are attributed to the same cluster. These algorithms can be used to find document relationships based on their metadata or full-text content. Fuzzy clustering techniques would be more suitable than hard clustering methods.
**Topic Maps:**

Topic Maps consist of a new standardized tool developed to model and manage knowledge structures and information resources. They are based on the concept of the traditional book index (list of information), but with the expanded capabilities of a thesaurus (interrelation of terms) and glossary (list of terms and definitions), all comprised into a simple and scalable model, which is very simple but effective. Knowledge “elements” (any set of subjects, ideas, concepts, themes) are classified into objects called topics, any kind of relationship among them can be made explicit in the form of associations, most of the topics can be addressed by one or more specific resources (occurrences), topics, associations and occurrences can be scoped for different contexts.

The main advantages of this technology are:

1. Its standardization provides flexibility and adaptability on its implementation.
2. The provision of an intuitive and meaningful (semantic) structure is translated into more efficient navigation and browsing processes among resources.
3. The ability to manage information resources and knowledge in the same tool.
4. And great scalability, as its architecture is designed to facilitate the merging of different topic maps without copying or modifying them.

One of the most obvious direct uses of TopicMaps is the navigation and browsing of information resources, although the combination of all its advantages makes them a very powerful tool for a wide range of information, knowledge and content management applications.

6.8 **Adaptive Knowledge-based Content Navigation**

In general, adaptive navigation systems enable personalized access to hyper-linked information. Adaptation in hypermedia systems can be provided at two different levels: at the presentation-level and at the link-level. Although both kinds of adaptation should be present in e-Learning systems that aim at providing flexible learning environments, link-level adaptation is particularly important to guide students through their learning path. Link-level adaptation deals with the discovery and display of relevant links to a given user. There are mainly two approaches for dynamically defining the links. One is to log the user’s actions so that the system can suggest
links based on past information. The other approach consists on keeping a record of the user’s current knowledge and interests in a profile and then search for pages that match the individual. The basic mechanism for adapting content to each student is based on the representation of both domain knowledge (domain model) and student knowledge level (student model) and uses the notion of pre-requisites and outcomes. Pre-requisites are basically the set of concepts a user needs to know to access a document and outcomes the set of concepts he/she is expected to acquire after reading it. The system then analyses the student’s knowledge – stored in the profile as weighted concepts to determine which links should be made available.

The knowledge domains are usually modeled manually by experts in the area. Instead of using such static models, fuzzy clusters and Topic Maps can be used for link-adaptation and for browsing the knowledge space. In [12], the authors also suggest the use of document clustering for adaptively linking resources. Their approach is to use hard clustering and adapt links in context of the user’s interests and of the documents’ contents. The advantage of having fuzzy clusters is that links can be ordered by degree of relevance, computed from the fuzzy memberships, and unobvious links may be revealed.

**Prototype**

A prototype of the knowledge representation framework has been implemented and the diagram shown in Figure 6.3 illustrates the complete process of knowledge discovery and representation:

---

![Knowledge Representation Framework Diagram](image)

**Fig 6.3 Knowledge Representation Framework**

Term weights were generated as a function of their frequency within the document and of their inverse document frequency (TF-IDF), i.e. the number of documents where the term occurred.
Then, the modified Fuzzy C-Means algorithm was applied to the TF-IDF data. The algorithm uses the cosine measure to assess the similarity between documents and iteratively updates the clusters centers until an objective function is optimized. After running the algorithm for a fixed number of clusters and subsequently a Topic Map was automatically generated.

Cluster-cluster relationships were computed as the cosine similarity between each pair of cluster prototypes, cluster-term relationships were obtained from the weight of the terms in the prototype vector and cluster-resource relationships consisted on the fuzzy memberships of each document in the cluster. Resource-resource relationships were computed as fuzzy relations and finally, term-resource relationships were derived in context of the clusters’ prototypes. The browser was used to visualize and navigate the Topic Map.

6.9 Automatic extraction of knowledge from the web:
Fig 6.4 Automatic Extraction of knowledge from web
Fig 6.5 Combined Knowledge Platform Functional overview

6.10 Instructional design and learning knowledge objects:

Design follows following arrangement.
Fig 6.6 Blooms taxonomy of knowledge verb representation.

SWRL (Semantic Web Rule Language) is used as the rule formalism. SWRL enables Horn-like rules to be used with an OWL knowledge base. SWRL rules can then exploits OWL classes, instances, and properties in their antecedent (body) and consequent (head).
• **Predefined methods**: exploit the knowledge already available in the organizational memory. For instance, if a theory requires the presentation of prerequisites, then there is a method able to use the domain ontology to retrieve the appropriate prerequisites. In fact, four predefined methods are offered for the moment:

• **The Learning Objective Method**: is used to find the learning objectives of the current learning knowledge object. These objectives can be found in the definition of the competence’s skills and in the competence description;

• **The Prerequisite method**: searches the appropriate prerequisites according to the competence ontology and to the learner model;

• **The Learner Score method**: is used to determine the actual score of a learner in a question or an exercise.

**Learning Knowledge Object**: as a learning object that is able to explain its own structure and the pedagogical intension behind it. As shown in the figure below.
The problem of creating learning materials can be solved either by using an authoring environment or by reusing existing resources. These resources can be either dedicated to training (pedagogical material) but can also take the form of domain documents such as reports or notes. In fact, working with pre-existing content is cost-efficient and communities of practice have a lot of electronic documents that can be re-used.

The research projects presented above focused more on training material. They focused on the generation of learning objects metadata and proposed ontologies to improve metadata, but neglected in general the use of ontologies to describe learning objects content except in (Gasevic et al., 2004) where domain ontologies are used to index learning object content.

The SeLeNe project (Keenoy et al., 2004) does not exploit domain ontology, whereas this research work enables to do this. The Trial Solution do use a thesaurus as its domain knowledge and tries to classify document content by searching the document for a list of sentences and keywords provided by the thesaurus.

This lightweight domain ontology does not reflect automatically new domain knowledge nor does it exploit document content. All the projects have the objective of creating a sort of learning
object repository whereas the objective here is to constituting a memory of knowledge objects (organizational memory) that can be used to dynamically assemble learning knowledge objects. Moreover, as far as the knowledge based integrated approach here natural language processing tools are used to represent document content into concept maps and to generate domain ontology.