CHAPTER 5

AN INTELLIGENT ONTOLOGY-BASED E-LEARNING TOOL FOR SOFTWARE RISK MANAGEMENT

5.1 INTRODUCTION

E-learning systems presently play a vital role as critical support mechanisms in educational institutions. However, the knowledge of these systems tends to be unstructured and static (Kontopoulos 2008). Through this work, a framework is proposed that uses an underlying ontology to organise course content based on semantics. Such an organization provides a meaningful structure to the courseware, and makes an effective understanding easy for the learner. The present systems (Lina and Rober 2010) also lack the ability to gain new knowledge during the course of their life-cycle, rendering them static. The proposed system aims to periodically update the knowledge presented by the system to keep track of current trends in the learning domain. Further, the system acts intelligent by initially assessing the newly-registered learner, in order to provide personalized suggestions based on a Bayesian network model. Emphasizing on the profound importance of risk management in the IT industry, the need of the hour is to educate potential software engineers; hence the learning domain of the proposed system is Software Risk Management. The proposed system thus focuses on providing an innovative tool to budding engineers, in order to equip them for the IT industry, and also intends to serve as a reference for professionals working on software projects.
5.2 SYSTEM ARCHITECTURE

The proposed system is developed as a web service, with the following four logical layers. As part of this research, a separate syllabus for software risk management is created and it is available in Appendix – II. Figure 5.1. shows the overall architecture of an intelligent ontology-based E-learning tool for software risk management.

![System Architecture Diagram](image)

Figure 5.1 System Architecture

5.2.1 Interface Layer

The interface layer contains the following three application-independent components:
5.2.1.1 Knowledge Diagnosis Tools

Knowledge diagnosis is performed on the newly registered learner to determine his/her knowledge level. The system presents the learner with a set of predetermined questions from each topic of Software Risk Management, which forms a node in the Bayesian network. The knowledge level of the learner in related topics is predicted, using the network model and presented to the learner as inferences.

5.2.1.2 Representation of Educational Resources

Java Server Pages are used to represent the educational resources in a learner-friendly manner. Each phase of the Software Risk Management is associated with an ontology whose visualization diagram is displayed, from which each topic may be selected for study.

5.2.1.3 Learner Profile Manager

The learner profile contains the personal information of the learner which is displayed on successful login. The scores secured by the learner on completion of the diagnosis test and computer-adaptive test at the end of the course, are updated in the profile and displayed to the learner.

5.2.2 Intelligent Layer

The intelligent layer consists of the intelligent components in the e-learning system.

5.2.2.1 Support for the Study Process

The subsystem for supporting the study process is the Bayesian network engine that produces inferences for the learner. Once the newly-
registered learner logs in to the system, his knowledge level in different topics of Software Risk Management is assessed using a diagnosis test. The results of this test are used to predict the knowledge level of the learner in related topics, and this indication is useful to the learner during the study process.

5.2.2.2 Search Logistics

The learner can search for a particular topic which may or may not be present in the local ontology. The subsystem of the search logistics initially searches for the requested topic in the local ontology. In case the information about the keyword requested by the learner is unavailable in the local ontology, the subsystem transfers the control to the module for accessing remote knowledge bases and updating the system knowledge accordingly. The system keeps track of all the topics searched by various learners. The ontology manager is informed of the list of topics that are not present in the local ontology to enable updation.

5.2.2.3 Subsystem of Intelligent Testing

The system is capable of testing the learner intelligently and generates questions dynamically on the basis of the performance of the learner in the previous question. Such computer-adaptive testing techniques are used widely in globally standardized tests such as GRE, TOEFL etc.

The computer-adaptive testing technique involves the following process. The learner is initially presented with a question of medium-level difficulty. If the question is answered right, then the learner is presented with a question of high-level difficulty. If not, the learner is presented with a question of low-level difficulty. Thus the testing subsystem adapts to the knowledge level of the learner. At the end of the test, the results are displayed and updated in the learner’s profile.
The dimensions of learning [100] is a comprehensive model used to define the learning process. The five dimensions of learning are

I. Attitude & Perception  
II. Acquire & Integrate Knowledge  
III. Extend & Refine Knowledge  
IV. Use Knowledge meaningfully  
V. Habits of Mind.

The diagrammatic representation of various dimensions of learning is given in Figure 5.2.

![Dimensions of Learning Diagram](image)

**Figure 5.2. Dimensions of Learning**

Among these five dimensions, this work mainly concentrates on ‘Use Knowledge Meaningfully’ dimensions, as the system has been developed with a specific objective of training the users who have been
recruited as a trainee and making the user more familiarize in software risk management domain. The system has also been developed with an optimistic view of the learner successfully attempting to answer all levels of difficulty and use this knowledge effectively during the software development process. The circle representing ‘Use Knowledge Meaningful’ subsumes the other two, and the circle representing ‘Extending & Refining Knowledge’ subsumes the circle representing ‘Acquiring and Integrating Knowledge’. Hence it is inferred that these two dimensions ‘Extending & Refining Knowledge’ and ‘Acquiring and Integrating Knowledge’ are implicitly presented in the learner if ‘Use Knowledge Meaningful’ is observed in a learner. The I and V dimensions namely ‘Attitude & Perception’ and ‘Habits of Mind’ are not considered in this work.

5.2.3 Learning Management System Layer

The Learning Management System layer contains the learners’ profiles and learning objects, consisting of the syllabus and courseware. The syllabus is structured, based on the underlying ontology to provide effective retrieval of semantically-related topics. Questions, classified on the basis of their difficulty, are also stored.

5.2.4 Deeper Knowledge Layer

The deeper knowledge layer contains the following components:

5.2.4.1 Local ontology

The local ontology contains information about the domain which is required for providing learning services. In the proposed system, the local ontology is termed SRMonto, which is an integration of ontologies developed for each phase of the Software Risk Management. The local ontology is integrated with the system by the developer, and periodically updated by the ontology manager.
5.2.4.2 Ontology Manager

The ontology manager keeps track of the various topics searched by the different learners, and the current trends in the learning domain. Topics that are not present in the local ontology may be updated periodically by the ontology manager.

5.2.4.3 Remote knowledge base

Remote knowledge bases may be similar to DBPedia, that contains structured information extracted from the Wikipedia.

5.3 MODULE DESCRIPTION

5.3.1 Learner Module

The new learner registers into the system by providing his/her personal details. The registered learner may login with his/her email id and password. The newly registered learner first takes the diagnosis test to determine his/her knowledge level. At the end of the test, the module updates the learner profile with the score.

The learner module consists of three main functions: register, login and manage ontology, as explained below.

5.3.1.1 Function: Register()

The register() function initially displays the registration form. The learner is prompted to fill the form and if any errors are present, they are indicated to the learner. On successful entry, the learner database is updated with the newly registered learner’s profile information. The algorithm for registration process is shown in Figure 5.3.
function register() returns success or failure
begin
    display_registration_form();
    retrieve_user_input();
    if errors in form then
        begin
            indicate_errors();
            return failure;
        end;
    else
        begin
            update database with new learner profile;
            display_diagnosis_test();
            return success;
        end;
    end;
end;

Figure 5.3. Algorithm for function register()

5.3.1.2 Function: Diagnosis_Test()

After the new learner registers into the system, the diagnosis test is displayed. A total of 20 have been identified from the key areas of software risk management such as Risk Factors, Categories of Risk, Implementation Factors, Risk Assessment, Planning Process, Control Process, Risk Identification Techniques, Qualitative Risk Analysis Techniques, Tree based Techniques, Techniques for Dynamic System, Risk Management Strategies and Risk Control Strategies. A Bayesian Network Model is constructed to predict the knowledge level of the newly-registered learner. The joint
probability distributions of the above identified related topics are calculated based on the similarity of topics.

Rule Syntax:
If \(<\text{topic1}> = \text{high}\) Then \(<\text{topic2}> = \text{high}\)

Rules of the above form are identified to form the Bayesian Network Model. The rule may be interpreted as “if the knowledge level of the learner in topic1 is determined to be high, then, based on the model, the knowledge level of the learner in topic2 may also be predicted as high.” The following rules have been identified for the domain Software Risk Management. The rules are visualized by the Figure 5.4.

RULE 1: If \(\text{risk\_factors} = \text{high}\) Then \(\text{categories\_of\_risk} = \text{high}\)

Rule 1 states that if the knowledge level of the learner in the topic “Risk Factors” is determined to be high, then it may be predicted that the knowledge level of the learner in the topic “Categories of Risk” is also high.

RULE 2: If \(\text{risk\_identification\_techniques} = \text{high}\) Then \(\text{qualitative\_risk\_analysis\_techniques} = \text{medium}\)

Rule 2 states that if the knowledge level of the learner in the topic “Risk Identification Techniques” is determined to be high, then it may be predicted that the knowledge level of the learner in the topic “Qualitative Risk Analysis Techniques” is medium.

RULE 3: If \(\text{implementation\_factors} = \text{high}\) Then \(\text{risk\_assessment} = \text{medium}\)

Rule 3 states that if the knowledge level of the learner in the topic “Implementation Factors” is determined to be high, then it may be predicted
that the knowledge level of the learner in the topic “Risk Assessment” is medium.

**RULE 4:** If risk_control_strategies = high Then risk_management_strategies = high

Rule 4 states that if the knowledge level of the learner in the topic “Risk Control Strategies” is determined to be high, then it may be predicted that the knowledge level of the learner in the topic “Risk Management Strategies” is also high.

**RULE 5:** If qualitative_risk_analysis_techniques = high Then tree_based_techniques = medium

Rule 5 states that if the knowledge level of the learner in the topic “Qualitative Risk Analysis Techniques” is determined to be high, then it may be predicted that the knowledge level of the learner in the topic “Tree-Based Techniques” is medium.

---

**Figure 5.4 Bayesian Network model for Software Risk Management**

where —— indicates “high implies high” relationship

——— —— indicates “high implies medium” relationship
RULE 6: If planning_process = high Then control_process = high

Rule 6 states that if the knowledge level of the learner in the topic “Planning Process” is determined to be high, then it may be predicted that the knowledge level of the learner in the topic “Control Process” is also high.

RULE 7: If techniques_for_dynamic_systems = high Then tree_based_techniques = medium

Rule 7 states that if the knowledge level of the learner in the topic “Techniques for Dynamic Systems” is determined to be high, then it may be predicted that the knowledge level of the learner in the topic “Tree Based Techniques” is medium.

RULE 8: If implementation_factors = high Then risk_factors = medium

Rule 8 states that if the knowledge level of the learner in the topic “Implementation Factors” is determined to be high, then it may be predicted that the knowledge level of the learner in the topic “Risk Factors” is medium.

COMPLEX RULE 1: If Qualitative Risk Analysis Techniques = high and Techniques for Dynamic System = high Then Tree Based Technique = Medium

It states that if the knowledge level of the learner in the topics “Qualitative Risk Analysis Techniques” and “Techniques for Dynamic Systems” are determined to be high, then it may be predicted that the knowledge level of the learner in the topic “Tree Based Techniques” is medium.

COMPLEX RULE 2: If Risk_Factors = Medium or Risk_Assessment = Medium Then Implementation_Factor = High.

It states that if the knowledge level of the learner in the topics “Risk Factor” or “Risk Assessment” is determined to be medium, then it may be predicted
that the knowledge level of the learner in the topic “Implementation Factor” is High.

Questions for the Diagnosis Test have been selected from topics that form the “If” clause of each rule in the Bayesian Network. The knowledge level of the learner in that particular topic is determined based on the number of questions that are answered rightly for that topic. This determined knowledge level is then applied to the model, and inferences about the knowledge level of the learner in the “Then” clause of each rule are made. These inferences are presented to the learner to indicate in the topics which they are deficient. Figure 5.5. shows the algorithm for function diagnosis_test().

```plaintext
function diagnosis_test() returns inferences
begin
  select_questions();
  display_questions();
  retrieve_learner_answers();
  determine_knowledge_level();
  predict_inferences();
  return inferences;
end;
```

**Figure 5.5 Algorithm for function diagnosis_test()**

The list of questions used for the diagnosis test is given below.

1. ________ organizes a large number of ideas into relationships.
   a. Assumption Analysis  b. Decision-making Analysis
   c. Affinity Grouping Technique  d. Checklists  **Ans: c**
2. When not to use Affinity Diagram?
   a. Information Retrieval should be easy
   b. When you are confronted with many facts or ideas in apparent chaos
   c. When issues seem too large and complex to grasp
   d. When group consensus is necessary  \textbf{Ans: a}

3. Which of these techniques can be used to rate the project status?
   a. Assumption Analysis
   b. Decision-making Analysis
   c. Affinity Grouping Technique
   d. Checklists  \textbf{Ans: d}

4. ______ refers to the adequacy with which the selected development model, process required for the specific program.
   a. Capability
   b. Suitability
   c. Familiarity
   d. Product Control  \textbf{Ans: b}

5. Which of the following is not a part of Development Process?
   a. Suitability
   b. Formality
   c. Usability
   d. Product Control  \textbf{Ans: c}

6. Which of the following refers to the adequacy with which the selected development model supports the specific program?
   a. Suitability
   b. Product Control
   c. Process Control
   d. Capacity  \textbf{Ans: a}

7. Which of these is an empirical approach used to determine the probability of occurrence of risk, based on the ratio of risk occurrence of the past projects?
   a. Classical Probability
   b. Frequency-based Probability
   c. Subjective Probability
   d. Objective Probability  \textbf{Ans: b}
8. Which of these is not a part of implementation factor in Risk Planning?
   a. Availability of staff  
   b. Environment  
   c. Funds  
   d. Probability of risk occurrence  
   **Ans:** b

9. The most preferred approach of risk control is _____
   a. Acceptance  
   b. Avoidance  
   c. Contingency  
   d. Incident Recovery Plan  
   **Ans:** b

10. ________ describes the degree which organization is willing to accept the risk as trade-off to the expense of applying controls
   a. Risk Control  
   b. Risk Awareness  
   c. Risk Planning  
   d. Risk Appetite  
   **Ans:** d

11. ________ is a systematic and structured examination of planned process to identify risks.
   a. HAZOP  
   b. Qualitative Technique  
   c. FMEA  
   d. Decision Analysis  
   **Ans:** a

12. ________ is a procedure to analyse the potential failure modes within the system.
   a. HAZOP  
   b. FMEA  
   c. FMECA  
   d. Assumption Analysis  
   **Ans:** b

13. Which of the following information is not included in the planning process?
   a. Recommended controls
   b. Lists of responsible teams and staff
   c. Start date for implementation
   d. Various risks identified  
   **Ans:** d
14. Which of the following process can be aided and expedited using risk plan?
   a. Risk Mitigation  
   b. Risk Avoidance 
   c. Risk Contingency  
   d. Risk Acceptance  
   Ans: a

15. ______ is a success-oriented system analysis method.
   a. GO method  
   b. Digraph method  
   c. Morkov Modelling  
   d. Dynamic Event Tree Analysis  
   Ans: a

16. ______ provides an integrated framework to explicitly treat time, process variables and system behaviour.
   a. DYLAM  
   b. FMECA  
   c. HAZOP  
   d. FMEA  
   Ans: a

17. ______ method is used for illustrating the sequence of outcomes which may arise after the occurrence of a selected initial event.
   a. Event Tree Analysis  
   b. Fault Tree Analysis  
   c. Dynamic Event Tree Analysis  
   d. Cause and Consequence Analysis  
   Ans: a

18. Which of these approaches are used to represent the operator behaviour?
   a. Event Tree Analysis 
   b. Fault Tree Analysis 
   c. Cause and Consequence Analysis 
   d. Dynamic Event Tree Analysis  
   Ans: d

19. ______ is a combination of fault tree analysis and event tree analysis.
   a. Management Oversight Risk Tree  
   b. Cause-Consequence Analysis  
   c. Dynamic Event Tree Analysis  
   d. Safety Management Organization Review Technique  
   Ans: b
20. _______ is a classical modeling technique used for assessing the time-dependent behaviour of many dynamic systems.
   a. Fault Tree Analysis                       b. Dynamic Event Tree Analysis
   c. Cause and Consequence Analysis  d. Markov Analysis       Ans: d

5.3.1.3 Function: Login()

The login() function allows the learner or the ontology manager to login to the system. If the learner wishes to login, he provides his email id and password, which are checked with the learner database to authenticate the learner. If the learner is new, the diagnosis test is displayed. For learners who have already taken the diagnosis test, the learner profile is displayed. The ontology manager may login to the system by providing a special login. On successful login, the homepage of the ontology manager will be displayed. Figure 5.6 shows the Algorithm for function login().

```
function login() returns success or failure
begin
  if email="admin" and password="adminadmin"
  begin
    display_OntologyManager_Home();
    manage_ontology();
  end;
  else
  begin
    search_database_for_profile();
    if found then
      begin
        if learner is new then
          display_diagnosis_test();
        else
          retrieve_learner_profile();
          return success;
        end;
      else
        return failure;
      end;
    else
      return failure;
  end;
end;
```

Figure 5.6 Algorithm for function login()
5.3.1.4 Function: Manage_Ontology()

The ontology manager may update the local ontology, based on the search patterns of various learners. The manage_ontology() function displays the list of keywords searched by various learners from which the ontology manager may determine the relevant topics that are not present in the local ontology, and update the information into the local ontology. Current trends in the learning domain may also be updated. Figure 5.7 shows the algorithm for function manage_ontology().

```
function manage_ontology() returns updated ontology
begin
    connect_to_Search_Table();
    retrieve_Table_Details();
    select_search_topics();
    update_local_ontology();
    search_current_trends();
    update_local_ontology();
    return updated ontology;
end;
```

**Figure 5.7. Algorithm for function manage_ontology()**

5.3.2 Course Material Module

The course material is represented by SRMONTO, the consolidated ontology for the five phases of Software Risk Management. Based on the suggestions provided by the system at the end of the diagnosis test, the learner may choose any of the five phases, which displays the ontology diagram, which depicts the class hierarchy in the form of a tree, with each concept being represented as a node. The learner may select any one of the leaf nodes to obtain information about it. Figure 5.8 shows the algorithm for displaying the course material.
function display_course_material(SRM_phase) begin
    display_ontology_diagram();
    retrieve_user_input();
    if selected class is in ontology
        display_syllabus_contents();
    end;

Figure 5.8 Algorithm for function display_course_material()

5.3.3 Feedback Module

A Computer-Adaptive Testing (CAT) is used in this module, where the system initially presents the learner with a question of medium-level difficulty. If the question is answered right, then the learner is presented with a question of high-level difficulty. If not, the system presents the learner with a question of low-level difficulty. Thus, the level of difficulty is changed continuously based on the performance. At the end of the test, the score is updated in the learner profile database. Figure 5.9 shows the algorithm for function feedback().

function feedback() returns score begin
    display_medium_level_difficulty_question();
    retrieve_user_input();
    if selected answer is right then
        display_high_level_difficulty_question();
    else
        display_low_level_difficulty_question();
    update_database_with_score();
    update_user_profile();
    return score;
end;

Figure 5.9 Algorithm for function feedback()
5.3.4 Search Module

The search() function takes the keyword list entered by the learner as the input. If the search queries are present in the local ontology, then the corresponding results are retrieved and displayed to the learner. If the local ontology does not contain the information, then the remote knowledge base is accessed to retrieve the information. Figure 5.10 shows the algorithm for function search().

```
function search(keyword list) returns search results
begin
    search_local_ontology();
    if results found then
        display_results();
    else
        begin
            access_remote_knowledge_base();
            update_system_knowledge();
            display_results();
        end;
    end;
end;
```

**Figure 5.10. Algorithm for function search()**

5.4 IMPLEMENTATION

The proposed system has been implemented as a web service using J2EE technology. The user interface has been developed using Java Servlets and Java Server Pages. The web server used is GlassFish v2.1 and Java DB (Derby) database is used for database access. The ontology SRMONTO has been developed using the Protégé software tool, and visualized using OWLViz plugin.
The home page displays an introduction about the e-learning tool as shown in Figure 5.11. The learner may click on the “Sign up” button which displays the registration page as shown in Figure 5.12.

Figure 5.11 Home Page

Figure 5.12 Registration Page
On successful registration, the learner may log in as shown in Figure 5.13. The system displays the diagnosis test for the new learner as shown in Figure 5.14 and the learner profile for the other learners is displayed as shown in Figure 5.15.

Figure 5.13 Learner Login Page

Figure 5.14 Learner Profile Page
Figure 5.15 Diagnosis Test Page

Figure 5.16 shows the results of the diagnosis test are displayed. The learner may then click on any one of the links provided in the sidebar, which displays the corresponding ontology visualization diagram as shown in Figure 5.17 and the related course material as shown in Figure 5.18 and .
The learner may choose to test himself after finishing the course. The system has been developed with an optimistic view of the learner successfully attempting to answer all levels of difficulty. The Figures 5.19,
shows a test results obtained after the successful completion of a user. It merely indicative score in the format of ratio between correctly answered questions (16) and total number of questions (30) provided.

Figure 5.19 Test Results Page

The learner can search for a topic that may be presented in the local ontology or accessed from the remote knowledge base as shown in Figure 5.20.

Figure 5.20 DBPedia Page for Turing machine
The ontology manager may login to the system periodically and the search patterns of various learners are displayed as shown in Figure 5.21. Every topic that has been searched from the remote knowledge base, if relevant to the domain, will be added to the local ontology, enabling the knowledge of the system to stay updated.

Figure 5.21 Search history page

5.5 SUMMARY

In this chapter, a novel method for learning software risk management, using an ontology based e-learning system has been proposed. The ontology developed for the system has been effectively represented by the OWL format, and it has also been represented semantically and visually. In order to eliminate the discrepancies, the curriculum for software risk management has been developed, based on the developed ontology. The system periodically updates the local ontology, based on the searching content of the user. In future, a discussion forum may be developed for enabling interaction between the learners. Updating the course material may be automated, based on changes in the ontology. An ontology based web service for software risk management is described in the next chapter.