



P-1447



KNOWLEDGE REPOSITORY SYSTEM

By

ASHEED VASUDEVAN
Reg. No. 71202621007



Of

KUMARAGURU COLLEGE OF TECHNOLOGY, COIMBATORE

A PROJECT REPORT

Submitted to the

FACULTY OF INFORMATION AND COMMUNICATION ENGINEERING

*In partial fulfillment of the requirements
for the award of the degree*

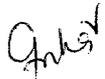
Of

MASTER OF COMPUTER APPLICATION

June, 2005

BONAFIDE CERTIFICATE

Certified that this project report titled **Knowledge Repository System** is the bonafide work of **Mr. Asheed Vasudevan** who carried out the research work under my supervision. Certified further, that to the best part of my knowledge the work reported herein does not form part of any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

**Faculty Guide****Head of the Department**

Submitted for the viva-voce examination held on _____

**Internal Examiner**
External Examiner 23/6/05

ABSTRACT

Knowledge Repository System(KRS) has been developed as an online knowledge repository Portal to store, organize, search and publish to the web any kind of documents: Word, Excel, PDF, HTML or PowerPoint.

KRS streamlines the entire documentation process for the Technical Experts to share information with each other and within the group. It is also customizable for any Technology and has complete integrated knowledge management of best practices, Case Studies, Tutorials and Training Materials. The product is made-up of independent modules developed in ASP. The product is architected into two modules as Repository module and Online Certification Mock Test module.

In the Repository Module portal provides role based access to its users to Upload documents, Upload TechTips, Upload Training Schedule, Upload AskUs questions and View to the documents, TechTips, Training Schedule according to the technology to which they are subscribed. In Online Certification Mock Test module, the users can take up test and also upload questions to the database according to the role of users accessing the module.

The product functionalities have been tuned for handling large amount of documents. It has been also created as a framework for all technologies for meeting the future needs.

ACKNOWLEDGEMENT

It is beyond the comprehension of mere elegance of word to acknowledge someone who has been the guiding spirit behind the dissertation. I am very much grateful to Dr. S.Thangasamy, Head of the Department, Department of Computer Science and Engineering for his excellent support.

With profound reference and high regards I record my indebtedness and gratitude to Mr. A.Muthukumar, Course Coordinator and to my guide Ms. V.Geetha, Senior Lecturer, Department of Computer Applications, Kumaraguru College of Technology for their guidance throughout my project work.

My heartfelt gratitude to Mr. Krishnan Srinivasan, Manager Projects, Data Warehousing and Business Intelligence Solutions Center, Cognizant Technology Solutions India Pvt Ltd, Chennai and Mr. Rengarajan Ramji, Senior Associate for their valuable guidance and their patience, without which this project would not have been completed. I take this opportunity to thank all the Associates in Cognizant Technology Solutions India Pvt Ltd, Chennai for their help and cooperation, which has led to the successful completion of this project work.

TABLE OF CONTENTS

CHAPTERS	PAGE NO.
ABSTRACT	iii
ACKNOWLEDGEMENT	iv
LIST OF FIGURES	vii
LIST OF TABLES	viii
LIST OF ABBREVIATIONS	ix
1. INTRODUCTION	1
1.1 COMPANY PROFILE	1
1.2 PROJECT OVERVIEW	3
1.2.1 Knowledge Repository – Introduction	3
1.2.2 KRS Highlights	4
1.2.3 KRS Components and Functionality	4
2. REQUIREMENT ANALYSIS	5
2.1 PROBLEM STATEMENT	5
2.2 FEASIBILITY ANALYSIS	6
2.2.1 Existing System	6
2.2.2 Proposed System	6
2.2.3 Technical Feasibility	7
2.2.4 Economic Feasibility	8
2.3 SOFTWARE REQUIREMENT SPECIFICATION	8
2.3.1 Purpose Of The Project	8
2.3.2 Scope Of The Project	9
2.3.3 Hardware Requirements	9
2.3.4 Software Requirements	10
2.3.5 Functional Requirements	11
2.4 SOFTWARE OVERVIEW	11
2.4.1 Active Server Pages	11

2.4.2 MS SQL Server	13
2.5 GENERAL CONSTRAINTS	14
3 SYSTEM DESIGN AND DEVELOPMENT	17
3.1 DESIGN CONCEPTS	18
3.2 MODULARIZATION	20
3.2.1 A Prototype Model for KRS	21
3.2.2 Document Management (Upload)	21
3.2.3 Document Management (Retrieve)	21
3.2.4 Online Certification Mock Test	21
3.2.5 E-Learning	22
3.3 DATA FLOW DIAGRAMS	22
3.3.1 DFD Level 0 - Context Flow Diagram	23
3.3.2 DFD Level 1	24
3.3.3 DFD Level 2 - Authenticate User	25
3.3.4 DFD Level 2 - Upload System	26
3.3.5 DFD Level 2 - Retrieve System	27
3.3.6 DFD Level 2 - Online Test	28
3.4 DATABASE DESIGN	29
3.5 DEVELOPMENT APPROACH	34
4 SYSTEM TESTING	35
4.1 WHITE BOX TESTING	35
4.2 BLACK BOX TESTING	36
4.3 ACCEPTANCE TESTING	36
4.4 ALPHA TESTING	36
4.5 BETA TESTING	37
5 SYTEM IMPLEMENTATION	38
5.1 IMPLEMENTATION PROCESS	38
5.2 FUTURE ENHANCEMENTS	39
6 CONCLUSION	40
APPENDICES	41
REFERENCES	54

LIST OF FIGURES

FIGURE NO	TITLE	PAGE NO.
3.1	Architecture Diagram	20
3.2	Data Flow Diagram Level 0	23
3.3	Data Flow Diagram Level 1	24
3.4	Data Flow Diagram Level 2	25
3.5	Data Flow Diagram Level 2	26
3.6	Data Flow Diagram Level 2	27
3.7	Data Flow Diagram Level 2	28
A 1.1	Home Page of the Portal	41
A 1.2	Screen to Upload Documents	42
A 1.3	Screen to Upload Training Schedule	43
A 1.4	Screen to Upload TechTips	44
A 1.5	Screen to View Documents	45
A 1.6	Screen to Manage Documents	46
A 1.7	Screen to Edit Subscriptions	47
A 1.8	Screen to Post Questions	48
A 1.9	Screen to view AskUs	49
A 1.10	Screen to Compile Online Test	50
A 1.11	Screen to Select Exam Category	51
A 1.12	Screen of Exam Page	52
A 1.13	Screen to Exam result	53

LIST OF TABLES

TABLE NO	TITLE	PAGE NO.
3.1	KR_DWSC_USER_MST	30
3.2	KR_USER_MST	30
3.3	KR_USER_TYPE_MST	30
3.4	KR_TECHNOLOGY_MST	30
3.5	KR_CLASSIFICATION_MST	31
3.6	KR_USER_SUBSCRIPTION_DTL	31
3.7	KR_DOCUMENT_ACCESS	31
3.8	KR_Tech_TIPS	31
3.9	KR_TRAINING_SCHEDULE	32
3.10	KR_UPLOADED_DOC_DTL	32
3.11	KR_QUESTION_DETAILS	32
3.12	KR_ANSWER_DETAILS	33
3.13	KR_TEST_RESULTS	33
3.14	KR_TRAINING_DETAILS	33
3.15	KR_CATEGORY_MST	33
3.16	KR_VERTICAL_MST	34
3.17	KR_VERTICAL_MASTER_DTL	34

LIST OF ABBREVIATIONS

ASP	-	Active Server Pages
HTML	-	Hyper Text Markup Language
IIS	-	Internet Information Server
KRS	-	Knowledge Repository System

CHAPTER 1

INTRODUCTION

This chapter is organized into three parts. The first part deals with the organization profile. It provides a brief insight into the history of the organization and the projects undertaken by them. The second part gives an introduction about the project.

1.1 COMPANY PROFILE– COGNIZANT TECHNOLOGY SOLUTIONS

Cognizant began its software services business in 1994 as the in-house technology development center for The Dun and Bradstreet Corporation (D&B) and its operating units. In 1996, the Company became a division of Cognizant Corporation, after the split-up of the D&B Corp. In June 1998, Cognizant Corporation was again spun-off into many independent companies, and Cognizant Technology Solutions was formed as a division of IMS Health, the world's leading provider of information solutions to the pharmaceutical and healthcare industries. In June 1998, the Company completed its initial public offering and got listed on the NASDAQ under the symbol CTSH. In November 2002, IMS Health divested its majority interest in Cognizant through a tax-free split-off, thereby making Cognizant a truly widely held company traded on the NASDAQ

Headquartered in Teaneck, New Jersey, Cognizant has twelve state-of-the-art software development centers. Ten offshore development centers in India are spread across Chennai, Bangalore, Kolkata, Pune and Hyderabad. Cognizant has customers in over 30 countries around the world and has plans to expand its presence to China and Eastern Europe. With more than 15,000 employees worldwide, Cognizant Technology Solutions Corporation is a global leader in IT services and has pioneered the pace setting 4th generation IT services model. Focused on delivering strategic information technology solutions that address the complex business needs of its clients, Cognizant provides applications management, development, integration, and re-engineering, infrastructure management and a number of related services such as enterprise consulting, technology architecture, program management and change management through its onsite/offshore outsourcing model.

Cognizant is ISO 9001:2000 certified, and assessed at SEI-CMMi Level 5 and P-CMM Level 5. It is also assessed against the BS 7799 standard, the internationally recognized certification for security and business continuity. It was recently ranked as the top information technology company in Forbes' 200 Best Small Companies in America and in Business Week's Hot Growth Companies.

Cognizant is one of the first companies offshore to be fully verticalized. It specializes in financial services (including banking, insurance, securities and portfolio management, and card processing), healthcare (including payers, providers and pharmaceutical companies), retail and hospitality, manufacturing and logistics, media and information defined services and telecom.

1.2 PROJECT OVERVIEW

1.2.1 Knowledge Repository System – Introduction

Knowledge Repository System has been developed as an online knowledge repository Portal to store, organize, search and publish to the web any kind of documents: Word, Excel, PDF, HTML or PowerPoint. KRS streamlines the entire documentation process for the Technical Experts to share information with each other and within the group. It is also customizable for any Technology and has complete integrated knowledge management of best practices, Case Studies, Tutorials and Training Materials. The product is made-up of independent modules developed in ASP. The product is architected into three modules as Repository module (Document Retrieval & Document Upload), E-Learning and Online Certification Mock Test module.

In the Repository Module portal provides role based access to its users to Upload documents, Upload TechTips, Upload Training Schedule, Upload AskUs questions and View to the documents, TechTips, Training Schedule according to the technology to which they are subscribed.

E-Learning module deals with the integration of E-Training products to the KRS and provides access to the users according to the user subscription.

In Online Certification Mock Test module, the users can take up test and also upload questions to the database according to the role of users accessing the module.

1.2.2 KRS Highlights

The following are the important features of KRS:

- Generic Framework.
- Simple & Pleasing Look and Feel.
- Role Based Access to Documents, Functionality.
- Subscription Based Access to various Technologies for the end user.
- Technology specific applications such as Training, Newsletter can be integrated easily.
- Online Test for Certification and Compilation of Questions for Online Test.
- Tech Tips & Training Schedule Details.
- Document and TechTips search.
- Vertical and Horizontal technologies based access.
- Bulk upload of questions from excel file.

1.2.3 KRS Components and Functionality

KRS is a complete knowledge repository portal developed using ASP and SQL Server 2000. The whole system is Modularized into four:

- Document Management(Upload)
- Document Management(Retrieve)
- Online Certification Mock Test
- E-Learning

The functionality of each module is discussed in detail in the later part of the report.

CHAPTER 2

REQUIREMENT ANALYSIS

A complete understanding of the requirements is essential for the success of software development. The software scope, initially established by the system engineer and refined during the project planning, is refined in detail. Models of the required data, information and control flow, and operational behavior are created. Alternative solutions are analyzed and allocated to various software elements. This chapter presents the problem statement as definitive statements to be solved. The feasibility study evaluates the viability of the project and presents the recommended strategy adopted for the development.

2.1 PROBLEM STATEMENT

The main objective of the portal is to allow the users to store, organize, search and publish to the web any kind of documents: Word, Excel, PDF, HTML or PowerPoint. The portal should be able to accurately handle the documents and also should give access to those according to the user roles. The user, according to the technology to which he has subscribed, can access the stored documents. The portal should also facilitate the user to take up online mock certification exams of the technology to which he has subscribed.

2.2 FEASIBILITY ANALYSIS

Feasibility is the measure of how beneficial or practical the development of information system will be to an organization. Once the problem is explained the feasibility study is to be done to test whether the product is achievable. The feasibility study describes the degree of the usefulness of the product to the organization. The feasibility study can be divided into four phases. They are as follows:

2.2.1 Existing System

The existing system for storing documents is done in terms of files in secondary storage and which is shared throughout to the users. The documents are neither sorted according to the category nor technology. The access to the share is not authenticated and there was no restriction in terms of user roles. Moreover the documents were scattered around many shared file locations in the secondary storage.

Limitations of Existing System

In the existing system, all the documents were maintained in terms of files in secondary storages, they were not easily accessible by the users. As the documents were not sorted according to the category, the entire search process was very complex.

2.2.2 Proposed System

The proposed system provides easy access to the users for the documents. The proposed system categorizes the documents and gives role-

based access to the users. All these storage retrieval are done in an efficient way.

Advantages of the Proposed System

The proposed system implements categorized document management and also role based access to these documents. As the proposed system is an online portal, the users can retrieve the documents globally from any location.

2.2.3 Technical Feasibility

Technical Feasibility is a measure of practicality of a specific technical solution and the availability of technical resource and expertise. This deals with the study of building within the pre-established cost and schedule bounds, the technology that exist to develop all elements of the system, systems reliability on proven technologies, the possibility of defining the interfaces, performance and functional aspects, analysis of technical resources, risk associated with the technologies. Feasibility study on quality of the elements of the system, system's external environment, and system communications is performed.

The tool can be developed with the existing technology. The individual modules are to be developed as ASP pages. SQL SERVER 2000 was chosen as the backend engine, because of the huge data volume to be handled. The Tool works effectively even for huge amount of data.

2.2.4 Economic Feasibility

Economic Feasibility is a measure of the cost-effectiveness of a project or solution. The System has been designed to work for any type of flat file(fixed width or delimited) and it manages voluminous data. Since the effort to develop the product was found to be feasible, the development presents a good investment for the organization. Hence the above system is economically feasible.

2.3 SOFTWARE REQUIREMENT SPECIFICATION (SRS)

The purpose of this SRS is to document the requirement of the system in a defined format based on the feasibility study. This is an elaboration of requirement analysis and must be validated and approved by the customer with respect to the customer needs. It specifies the required behavior of a system in terms of input data, required processing, output data, operational scenarios and interfaces. The development environment and system perspective give the view of the system. It specifies the attribute of the systems including performance, security, maintainability, reliability, audit ability, availability and safety requirements and design constraints.

2.3.1 Purpose Of The Project

In today's global business environment, it is crucial for organizations to manage the documents related to the various technologies. Knowledge Repository System provides a complete knowledge, tech tips, training schedule etc related to the technologies to which the user has subscribed. These include uploading training schedule, post questions, upload techtips, upload documents and also upload questions for the online mock test. The online mock test

provides users to have an hands on experience to the model of certification exams for various technologies.

2.3.2 Scope Of The Project

The scope of the project is to convert requirements specified by the user into functional requirements and implement the same in the system. It involves the following stages Requirement Analysis, Functional Specification, Design, Coding, and Testing.

2.3.3 Hardware Requirements

Application Server:

Processor	:	Intel Pentium III
Speed	:	600MHz
Memory	:	128MB RAM
Hard Disk	:	10GB

Database Server:

Processor	:	Intel Pentium III
Speed	:	400MHz
Memory	:	128MB RAM
Hard Disk	:	80GB

Client Machine:

Processor : Intel
Speed : X86 Series
Memory : 32MB RAM
Hard Disk : 2GB

2.3.4 Software Requirements**Application Server:**

Operating System : Windows 2000 and Above.
Web Server : Internet Information Server.
Front-End : ASP and HTML.

Database Server:

Operating System : Windows 2000 and Above.
Database : MS SQL Server 2000.

Client Machine:

Operating System : Windows 9x and Above.
Software : Web Browser.

2.3.5 Functional Requirements

- The application should have capabilities like
 - Upload documents, techtips, training schedule, questions for exams and also post Frequently Asked Questions(FAQ's).
 - Role based access to the documents uploaded to the database, FAQ's, TechTips and Training Schedule according to the technologies to which they have subscribed.
 - Bulk upload of questions from excel file.
 - Online Certification Mock Test for various technologies.
 - Search TechTips and Documents
- The application must be able to handle voluminous documents.

2.4 SOFTWARE OVERVIEW

2.4.1 Active Server Pages



The first attempt at providing user interaction in web pages was through server-side scripts. The client browser would send a request to the server. A script on the server would process the request and send the results back to the client browser.

There are several methods in which we create Server based scripts. A popular option available to the web application developers today is Active Server Pages. Active Server Pages(ASP) is a Server-Side scripting environment, developed by Microsoft. ASP is a technology that provides a web server with a capability to process application logic and return standard HTML to the browser.

An ASP document has an .asp file extension and can have a combination of HTML, Client-Side Script, Server-Side Script, ASP objects, Applets, ActiveX components. When an ASP file is displayed in the browser, the HTML syntax is executed and the page is displayed in the browser. The Server based script is executed on the server and the resulting HTML is sent to the browser.

Advantages of using ASP:

- As a developer, we can make changes to the .asp file on the server and save the changes to the file. The next time the page is loaded, the script will be automatically compiled. The reason being, the ASP technology is built into all Microsoft Web Servers.
- We can create a web page with dynamic content. The web server processes the code and generates HTML depending on the request made by the user. For example we can display an Active Server Page that displays a greeting based on the time of day.
- By default, an ASP sends only ASCII text to the client browser. Typically, any browser running on any operating system can access the application embedded in the ASP scripts.

The user will be able to view the results of the script but cannot see the source code. That is because the script commands that generate the page are retained on the web server and not sent to the web browser. The user cannot make any changes to the script and thereby tamper with data or content of the web page.

2.4.2 MS SQL Server

Microsoft SQL Server is a Structured Query Language (SQL) based, client/server relational database. Each of these terms describes a fundamental part of the architecture of SQL Server.

Database

A database is similar to a data file in that it is a storage place for data. Like a data file, a database does not present information directly to a user; the user runs an application that accesses data from the database and presents it to the user in an understandable format.

Database systems are more powerful than data files. The data is more highly organized. In a well-designed database, there are no duplicate pieces of data that the user or application has to update at the same time. Related pieces of data are grouped together in a single structure or record, and relationships can be defined between these structures and records.

When working with data files, an application must be coded to work with the specific structure of each data file. In contrast, a database contains a catalog that applications use to determine, without being tied to a specific data format.

A database typically has two components: the files holding the physical database and the database management system (DBMS) software that applications use to access data. The DBMS is responsible for enforcing the database structure, including:

- Maintaining the relationships between data in the database.
- Ensuring that data is stored correctly and that the rules defining data relationships are not violated.
- Recovering all data to a point of known consistency in case of system failures.

Relational Database

There are different ways to organize data in a database but relational databases are one of the most effective. Relational database systems are an application of mathematical set theory to the problem of effectively organizing data. In a relational database, data is collected into tables.

Structured Query Language (SQL)

To work with data in a database, you must use a set of commands and statements (language) defined by the DBMS software. There are several different languages that can be used with relational databases; the most common is SQL. Standards for SQL have been defined by both the American National Standards Institute (ANSI) and the International Standards Organization (ISO). Most modern DBMS products support the Entry Level of SQL-92, the latest SQL standard.

2.5 General Constraints

The Various terms involved in Software performance requirements are briefly depicted below.

- Speed

Speed corresponds to the overall performance improvement of the software. By improving the data handling operations and providing the result to the user, the overall speed of the software should be maximum in order to minimize the response time.

- Quality of Service

The software is expected to provide excellent quality of service to the end user. By doing so the services provided by the organization, which the software serve, will largely benefit in terms of improvement in the quality.

- Availability

The availability of the software is very crucial and the software is expected to have very minimal downtime. Special considerations are required to improve the stability of the software and to reduce the operational downtime of the software.

- Reliability

The operations performed by the application and the results are expected to be precise. All the data handled by the software should be carefully handled so that the software will become reliable and produce accurate result.

- Expandability

Good software presents provisions for accommodation of the features that might be required and included in the future. Such provisions are expected to be a part of the software system design and implementation.

- Adaptability

The developed software should be an apt solution for the developed purposes. The software should be adaptive i.e., should include options that could be customized according to the requirements.

This system is based on the current procedures followed by Cognizant. Provisions have been made to accommodate any change in the management policies or in any procedures; the software team who are maintaining this system can make changes accordingly.

CHAPTER 3

SYSTEM DESIGN AND DEVELOPMENT

This is the pivotal point in the system development life cycle. User requirements have been identified. Information has been gathered to verify the problem and evaluate the existing system. A feasibility study has been conducted to review alternative solutions and provide cost/benefit justification. The culmination of the study is a proposal summarizing the findings and recommending a candidate system for the user. It translates the system requirements into an operational system.

Since the reason behind the candidate system made sense, the system is now ready to undergo the design phase. The design as Award puts, is *“a solution, a ‘how to’ approach, compared to analysis, a ‘what is’ orientation”*.

System design is the process of planning a new system or to replace or the complement of the existing system. The design is based on the limitations of the existing system and the requirement specifications that are gathered in the phase of system analysis.

System design primarily involves a logical design and then the physical construction of the system. During the logical design of the system, the detailed specifications for the new system have been provided by describing its features, procedures that meets the system requirements.

The requirement of the data was identified after the identification of the objectives. After identifying the objectives, performing the code design of the input documents, output formats and the frequency of reports are finalized.

3.1 DESIGN CONCEPTS

Several concepts such as abstraction, refinement, modularity, software architecture, control hierarchy, structural partitioning, data structures, software procedure, information hiding are kept in mind during the design phase. Each concept provides the designer with the formulation from which more sophisticated design methods can be applied.

- Modularity (both in program and data) and the concept of abstraction enable the designer to simplify and reuse the software components.
- Refinement provides a mechanism for representing successive layers of functional detail.
- Programs and data structure contributes to an overall view of software architecture, while procedure provides the details necessary for algorithm implementation.
- Information hiding and functional independence provide heuristics for achieving effective modularity.

Design principles

Design process is a set of iterative steps, which enable the designer to describe all aspects of the software to build. The main principles in software design are

- The design process should not suffer from 'tunnel vision'.
- The design should be traceable to the analysis model.
- The design should not reinvent the wheel.

- The design should minimize the intellectual distance between the software and the problem, as it exists in the real world.
- The design should exhibit uniformity and integration.
- The design should be structured to accommodate changes.
- Design is not coding, coding is not design.
- The design should be reviewed to minimize the conceptual Errors.

Architectural design

Architectural design involves identification of software components, decoupling and decomposing them into processing modules, conceptual data structures and specifying relationship among the components.

The Knowledge Repository system has been developed as a three-tier architecture application. The 3 tiers or layers are:

1) View Layer

This layer contains the HTML files, Active Server Pages, Scripting languages (JavaScript, VBScript) and style sheets which are used for designing the GUI for the Knowledge Repository system.

2) Middle layer

This layer is responsible for connecting to the database and processing the data in the database. The Knowledge Repository system uses IIS to process the requests.

3) Data Layer

This layer contains the database which stores the information required for the application. In the Knowledge Repository System, SQL Server database is the data layer.

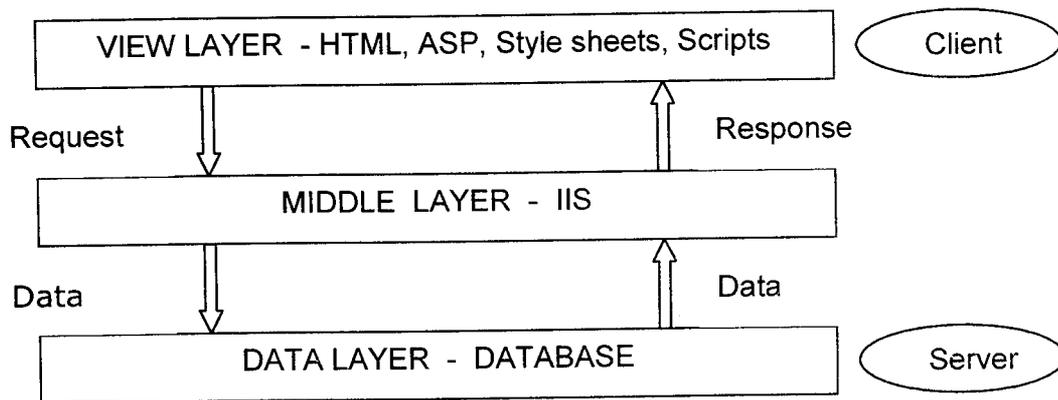


Figure 3.1: Three-tier Architecture layers

3.2 MODULARIZATION

Modularity is the single attribute of software that allows a program to be intellectually manageable. A reader cannot easily grasp monolithic software. The five criteria that enable us to evaluate a design method with respect to its ability to define effective modular design are Modular Decomposability which provides a systematic mechanism for decomposing, Modular compos ability which enables existing design components to be assembled into a new system, Modular understandability which enables us to understand a module as a single unit, Modular Continuity which forces the developers to make changes to individual modules rather than to entire system, and Modular Protection, which does not allow the aberrant conditions occurring in the module to affect the entire system. All these factors are considered in decomposing this System into modules that are integrated to satisfy the problem requirements.

3.2.1 A PROTOTYPE MODEL FOR KRS

A prototype model for Knowledge Repository System was developed for one technology. The files related to that technology alone was uploaded. The documents were uploaded and users were given access to them without any restriction. The documents were categorized under this technology.

3.2.2 DOCUMENT MANAGEMENT (Upload)

This part deals with the uploading of documents corresponding to the technologies to which the user has subscribed. Uploading of Training schedules, TechTip's, Questions for online mock test and post FAQ's are also dealt in this section.

3.2.3 DOCUMENT MANAGEMENT (RETRIEVE)

This module deals with the retrieval of documents from the database and display the document list according to the classification to which it falls. The documents displayed are only according to the technology to which the user has subscribed. Paging has also implemented for the listing of documents by confining to the restriction of ten documents per page. The retrieval of TechTips and Training Schedule are also dealt in this module and are displayed in the home page according to the technology to which the user has subscribed.

3.2.4 ONLINE CERTIFIATION MOCK TEST

This module deals with compilation of questions for the mock test and the Online Certification test pages. TechMasters can use 'Compile Online test' page to upload questions to the database. Users are also allowed to take up

the online mock test corresponding to the technology to which he has subscribed.

The questions are sorted and stored according to the technology and category under which it falls. If the category corresponding to a technology has more than 25 questions then the user is listed with that category name from which he can select the category and take up the test.

3.2.5 E-LEARNING

This module deals with the integration of e-learner corresponding to a technology by suppressing the authentication provided by it. The windows authentication username and password is used to create a user for that e-learner and allows the user to directly access the e-learner pages on just clicking on the e-learner link without any further authentication.

3.3 DATAFLOW DIAGRAMS (DFD)

Data Flow Diagram is a graphical representation that depicts information flow and the transforms that are applied as data move from input to output. The DFD can be used to represent a system or software at any level of abstraction. DFDs can be partitioned into levels that represent increasing information low and functional detail. For this system the DFDs are given up to Level 2 [Figure 3.7].

3.3.1 DFD LEVEL 0 – CONTEXT FLOW DIAGRAM

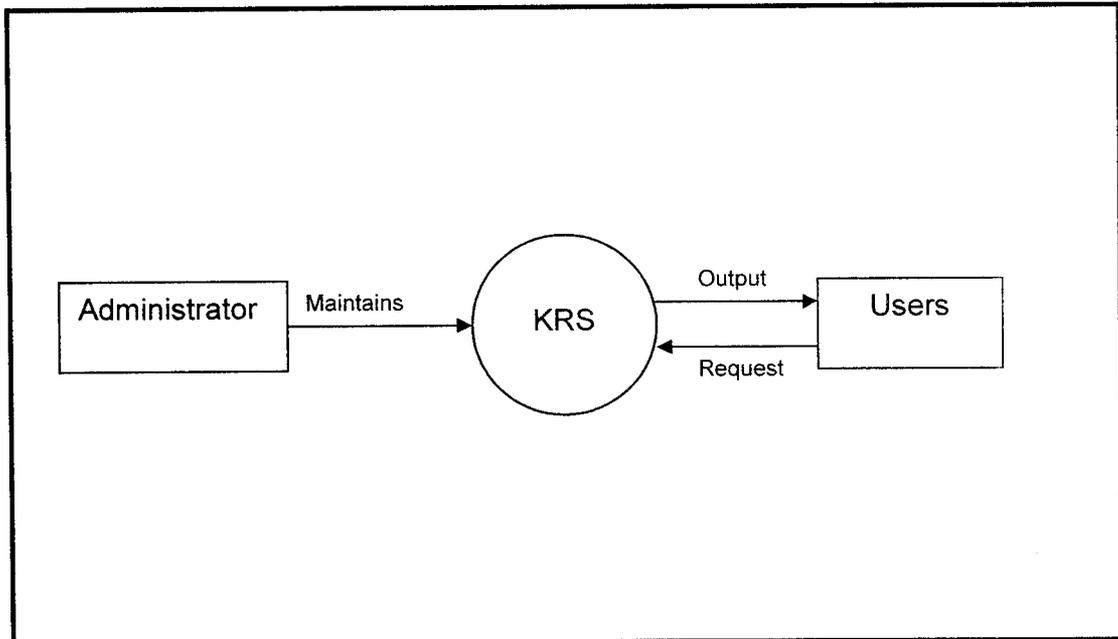


Figure 3.1: DFD LEVEL 0

3.3.2 DFD LEVEL 1

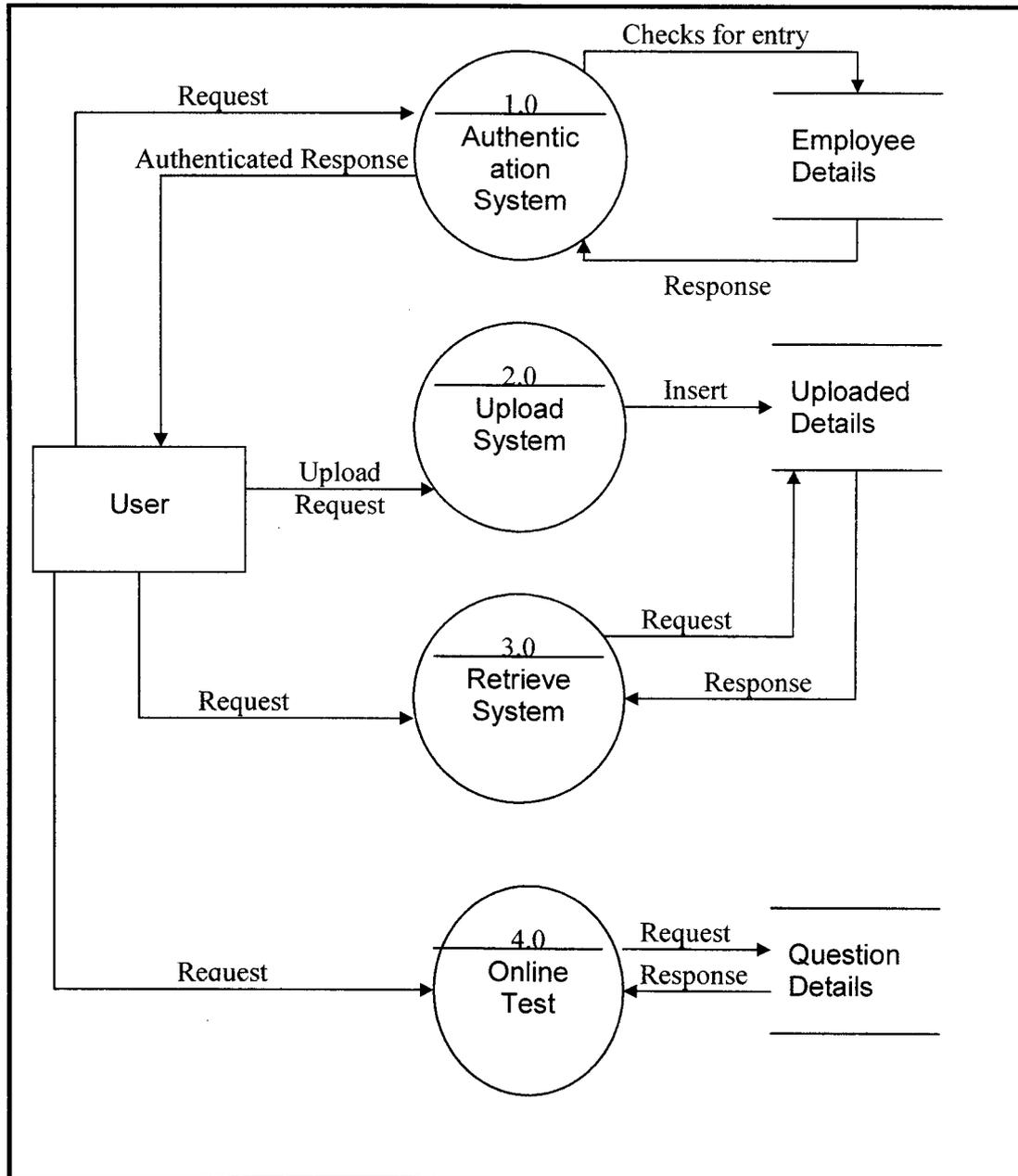


Figure 3.2: DFD LEVEL 1

3.3.3 DFD LEVEL 2 – AUTHENTICATE USER

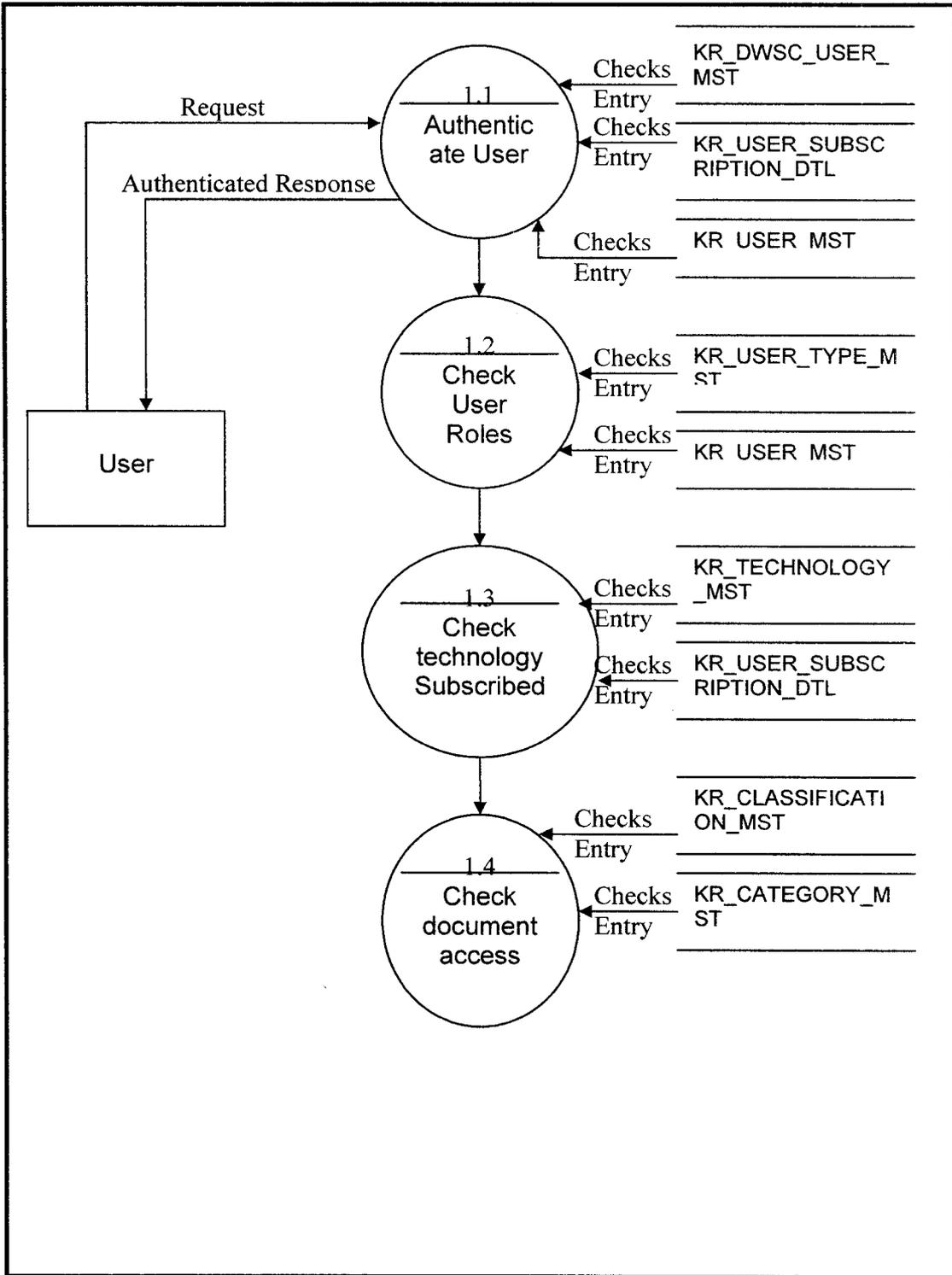


Figure 3.2: DFD LEVEL 2 - AUTHENTICATE USER

3.3.4 DFD LEVEL 2 – UPLOAD SYSTEM

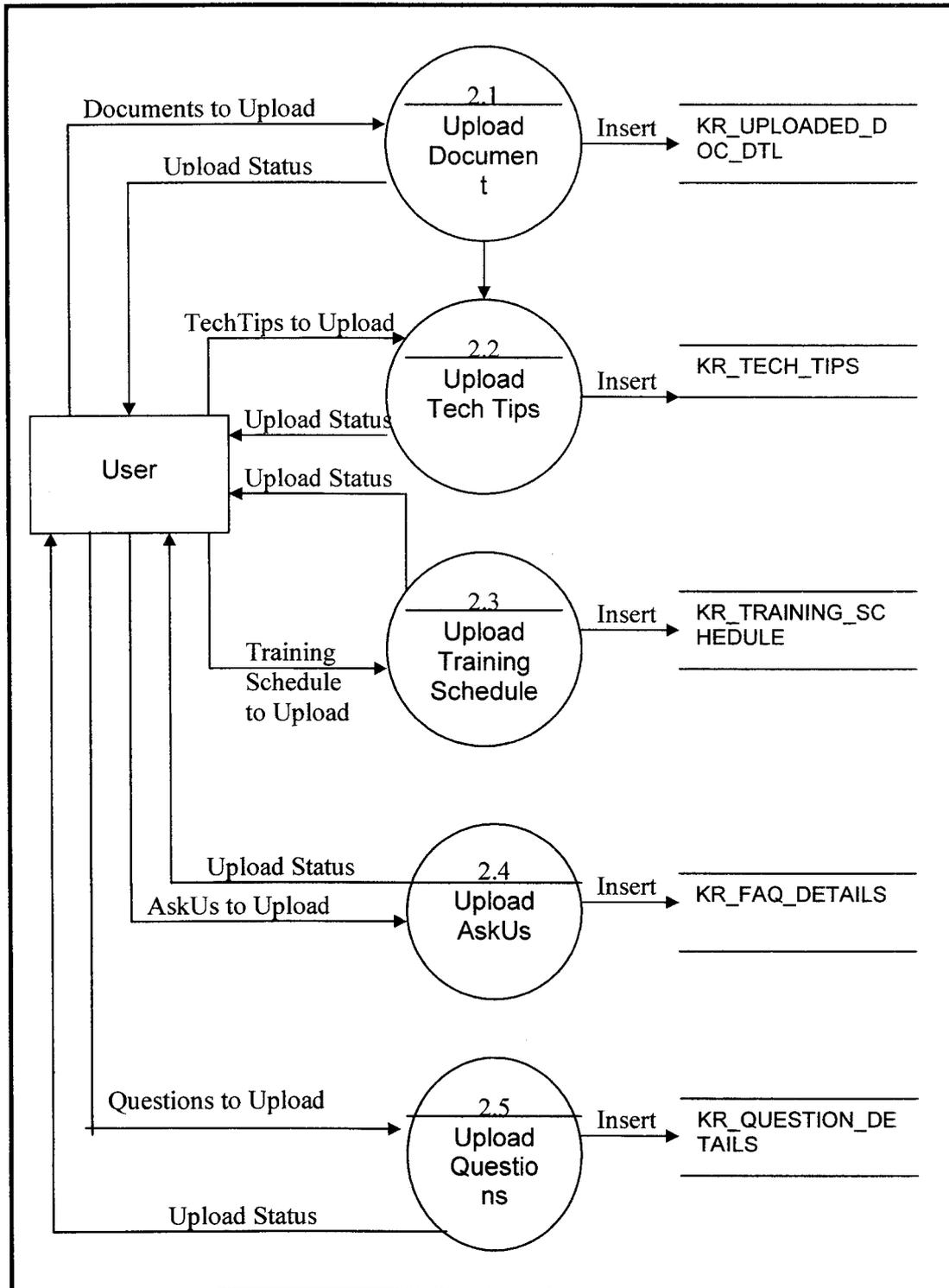


Figure 3.3: DFD LEVEL 2 - UPLOAD SYSTEM

3.3.5 DFD LEVEL 2 – RETRIEVE SYSTEM

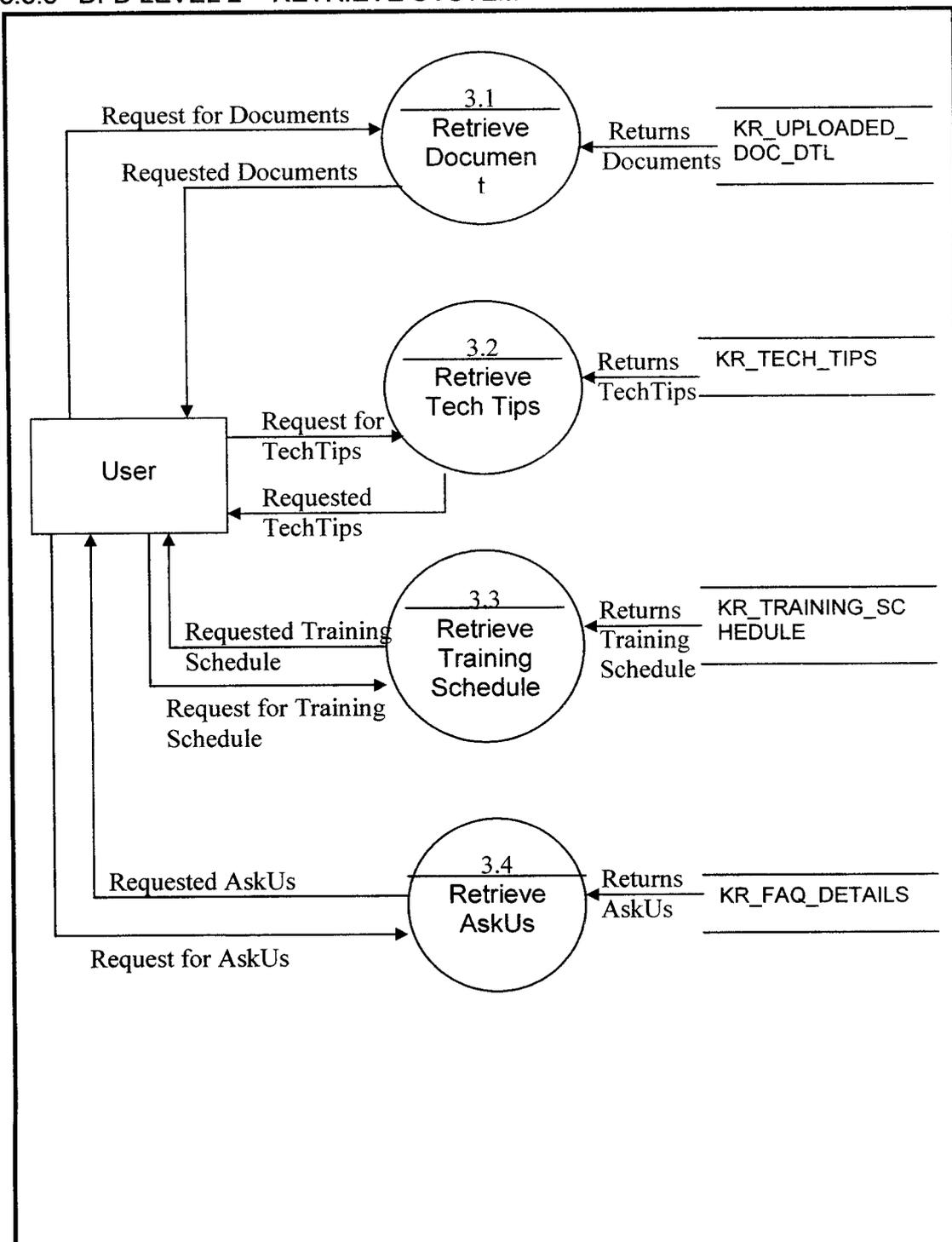


Figure 3.4: DFD LEVEL 2 - RETRIEVE SYSTEM

3.3.6 DFD LEVEL 2 – ONLINE TEST

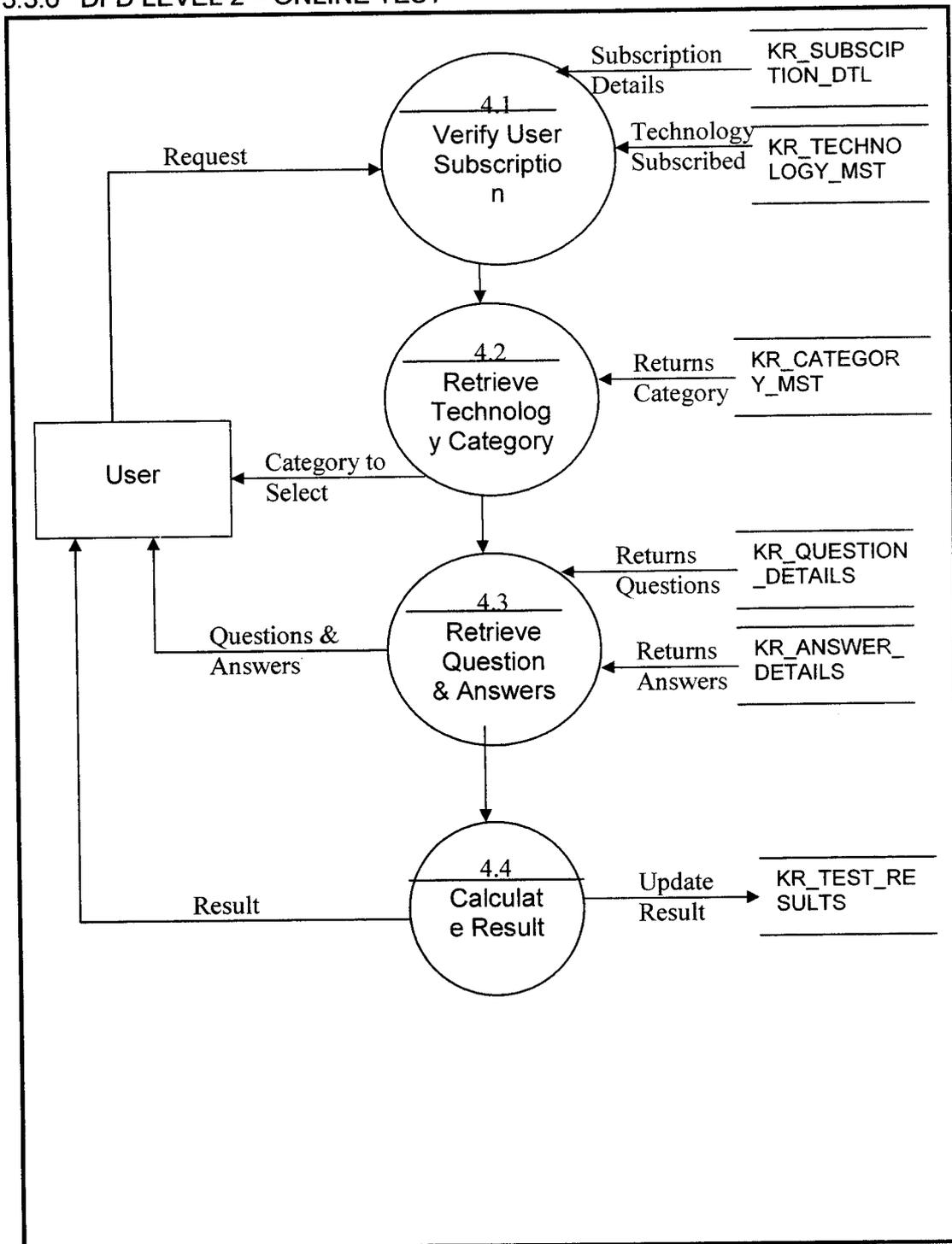


Figure 3.5: DFD LEVEL 2 - ONLINE TEST

3.4 DATABASE DESIGN

The general theme behind a database is to handle information as an integrated whole. A database is a collection of inter-related data stored with minimum redundancy to serve many users quickly and efficiently. The general objective is to make information necessary, quick, inexpensive and flexible for the user. The organization of data in a database aims to achieve 3 main objectives:

- Data Integration
- Data Independence
- Data Integrity

Normalization: Data structure is refined through a process called normalization. Data are grouped in the simplest way possible so that changes can be made with a minimum of impact on the data structure. When too many attributes are grouped together to form entities, some attributes are found to be entities themselves. Further normalization of these entities into attributes linked by common data elements to form relationships improves the effectiveness of the DBMS. The Knowledge Repository System uses SQL Server database to store and handle information related to the proposed system.

Table schemas for KRS

Table 3.1 - KR_DWSC_USER_MST

Description : Stores data related to the employees in data warehousing team.

Associate_ID	char
Associate_Name	char
Email_ID	char

Table 3.2 - KR_USER_MST

Description : Stores data about the users subscribed.

Associate_ID	varchar
Associate_Name	varchar
Email_ID	varchar
User_Type	varchar
Subscribed_Time	datetime

Table 3.3 - KR_USER_TYPE_MST

Description : Stores data related to the type of users.

User_Type	int
Description	varchar

Table 3.4 - KR_TECHNOLOGY_MST

Description : Stores data related to the different technologies.

Tech_id	int
Technology	varchar
URL	varchar
eTrainer	char
NewsLetter	char

Table 3.5 - KR_CLASSIFICATION_MST

Description : Stores data about the different classifications.

Classification_ID	int
Classification_Desc	varchar

Table 3.6 - KR_USER_SUBSCRIPTION_DTL

Description : Stores data about the technologies subscribed by the users.

Associate_ID	varchar
Tech_ID	int
isTechMaster	char

Table 3.7 - KR_DOCUMENT_ACCESS

Description : Stores data related to the type of documents the users can access.

User_Type	int
Classification_ID	int

Table 3.8 - KR_TECH_TIPS

Description : Stores data about the techtips uploaded.

Tips	varchar
Tech_ID	int
Eff_From_Date	smalldatetime
Eff_To_Date	smalldatetime

Table 3.9 - KR_TRAINING_SCHEDULE

Description : Stores data about the training schedule uploaded.

Training_Topic	varchar
Tech_ID	int
Trainer_Name	varchar
Date	smalldatetime
Venue	varchar
Duration	int
Training_Type	char

Table 3.10 - KR_UPLOADED_DOC_DTL

Description : Stores data about the documents uploaded.

File_Name	varchar
File_Description	varchar
File_Type	varchar
Physical_Location	varchar
Uploaded_By	varchar
Uploaded_Timestamp	smalldatetime
File_Search_Metadata	varchar
Tech_ID	int
Classification_ID	int

Table 3.11 - KR_QUESTION_DETAILS

Description : Stores data about the questions uploaded.

QNo	int
Question	varchar
QType	int
Tech_ID	int
Category_ID	int

Table 3.12 - KR_ANSWER_DETAILS

Description : Stores the answers corresponding to the questions uploaded.

QNo	int
QOption	varchar
QAnswer	int

Table 3.13 - KR_TEST_RESULTS

Description : Stores data related to the test results.

Login_ID	varchar
Attempt	int
Test_Date	smalldatetime
TimeStamp	smalldatetime
Score	int
Result	varchar
Tech_ID	int

Table 3.14 - KR_TRAINING_DETAILS

Description : Stores data related to the training details.

Certification_Topic	varchar
Certification_Description	varchar
Tech_ID	int

Table 3.15 - KR_CATEGORY_MST

Description : Stores data related to the categories under the technologies.

Category_ID	int
Tech_ID	int
Category_Name	varchar

Table 3.16 - KR_VERTICAL_MST

Description : Stores data related to the different verticals.

Vertical_ID	int
Vertical	varchar

Table 3.17 - KR_VERTICAL_MASTER_DTL

Description : Stores data related to the vertical Masters.

Associate_ID	int
Vertical_ID	varchar

3.5 DEVELOPMENT APPROACH

The two most commonly used approaches for the development and testing of software programs are: The traditional "Bottom-Up" approach and the contemporary "Top-down" approach. The Knowledge Repository system was developed using the latter approach. The Top-down approach involves the identification of a model of the organization's information needs and then the design of subsystems is based on this model.

The complete document sharing the report of the development phase was sent to the management for their approval. The reports of the development phase will be reviewed by the management team and any restrictions are then rectified.

CHAPTER 4

SYSTEM TESTING

System Testing is most vital activity that has to be enforced in any system development. This could be done parallel during the development phase and after implementation. The feedback received from this testing was carefully examined for further enhancements.

It is the part of testing where the entire website is tested. This testing is performed with the requirements document as the reference and the goal is to see whether the application meets the requirements

4.1 White Box Testing

White box testing also referred to as glass-box software testing. It is a test case design that would use the 'program control flow' structures to derive software test cases. The software engineer can derive white box software testing using following guidelines.

- All independent paths within a module have been exercised.
- All loops are executed at their boundaries and within their operational bounds.
- All internal data structures are exercised to assure their validity

All the independent paths in the four modules of KRS were tested and found consistent

4.2 Black Box Testing

Black box testing enables the software engineer to derive set of conditions that will fully exercise all requirements for a web design code. Also note that black box testing “is not an alternative” to white box testing. Rather it is a compulsory approach that it is likely to uncover a different CLASS OF ERRORS than what white box testing methods reveal. Black box testing reveal attempts to reveal errors in the following software work areas:

- Interface of inputs
- Database access
- Initialization and termination

The interfaces and all links in KRS were tested to find whether relevant pages for shown accordingly. The database connectivity and data retrieval were tested and found consistent.

4.3 Acceptance Testing

Acceptance testing involves planning and execution of the functional tests, performance tests and stress tests in order to demonstrate that the implemented system satisfies its requirements. It is not unusual for two sets of acceptance tests to be run those developed by the Quality Assurance group and those developed by the customer. The KRS was tested for load handling and data handling and found successful in handling large amounts of data and users.

4.4 Alpha Testing

A third person who just has the knowledge and the working capacity of the system conducts the alpha test at the developer’s site. The developer ‘looks over the shoulder’ of the user and records the errors and usage problems. The

user in turn gives general discomforts, which may be mended to make the system little better in a way of efficiency and user-friendly.

The KRS was handed over to a user and was asked to navigate through the system and try all possibilities and all the requirements were satisfied.

4.5 Beta Testing

After alpha testing is done the developed website is given to other solution partners to check for errors. After this testing has been done then the website will be published.

The KRS was undergone beta testing and found error free and was published.

CHAPTER 5

SYSTEM IMPLEMENTATION

Implementation includes all those activities that take place to creating a domain name, uploading the files to the website and other administrative activities. A proper implementation is essential to provide a reliable system to meet the requirements of the customers and the company. An improper installation may affect the success of the website and the administrative system.

5.1 Implementation Process

The implementation plan includes a description of all the activities that must occur to launch the portal and to put it into operation. It identifies the personnel responsible for the activities and prepares a time chart for launching the portal. The implementation plan consists of the following steps.

- Lists all files that have to be uploaded.
- Identifies all data required to build new files during the upload.
- Lists all new documents and procedures that go into the new system.

The implementation plan should anticipate possible problems and must be able to deal with them. The usual problems may be missing documents; mixed data formats between current files, errors in data translation, missing data etc. The implementation started after all the modules of the portal are integrated and tested together.

5.2 Future Enhancements

Enhancement is always an important and necessary activity in the life of the portal developed. Any system developed should always provide room to accept any changes and further inclusions.

This may occur so as to suit the new trend and additional facilities required by the web user. The future enhancements are to be carried out in such a way that it should not affect existing facilities in the websites adds more attractive features in the website.

The future enhancements involves the following features

- Facilities for uploading multiple files at a time.
- Features will be added such that whenever a new training schedule is uploaded, it will be intimated to the subscribers through automated mail.
- Facilities to upload multiple training schedule at a time.
- Upload capability of upto 50MB.

CHAPTER 6

CONCLUSION

By taking up this project for Cognizant Technology Solutions titled as Knowledge Repository System I attained an extensive knowledge of how to use the existing technologies to create a product based on the clients requirements. This project is aimed to provide assistance for the employees of Cognizant Technology Solutions to satisfy their knowledge search. The programming techniques used in the design of the system provide a scope for further expansion and implementation of any changes, which may occur in future. Maximum care and concentration has been focused to troubleshoot this project. This design can be applied for any organization under similar circumstance.

Invaluable experience has been gained in the areas of users system and database design, system testing and implementation. Since the requirements of any organization and their standards are changing day to day the system has been designed in such a way that its scope and boundaries could be expanded in future with little modification. The portal has been thoroughly tested and all the client needs are assured. Key issues in security has been checked for consistency. The tool has been successfully launched by giving rights to authorized users. Users can gain access to the portal by subscribing to the technology of his choice. This project has been taken up with at most dedication and it has been developed in such a way to accommodate any type of changes.

APPENDIX 1

SNAPSHOT OF SCREENS

1. Repository Phase

**DATA WAREHOUSING
KNOWLEDGE REPOSITORY**

Home Documents Certification eTrainer NewsLetter Upload Tech Tips Training Schedule Manage Documents Ask Us My Subscriptions Feedback/HelpDesk

MicroStrategy and Cognizant: Best Practices for BI

Welcome to the **MicroStrategy, Teradata, Informatica, Cognos, Business Objects & Data Stage Knowledge Repository**, your world-class teaching and learning resource on the Technology you have subscribed. Here you can find materials such as Manuals, Tech Notes, Case Studies, Proposals from the respective Technology experts.

[SDK for Web Developer Guide-Web Application Development.pdf](#)
[TCP Program Certification Guide.pdf](#)
[Introduction to Microstrategy71.pdf](#)

If there are multiple files to be loaded into an Empty target table then you can do it using FastLoad script without END LOADING command till the last file. After the last file has been loaded you need to run an END LOADING command.

MicroStrategy Basics
 This course helps to learn how to design a logical data model and how to translate a logical data model into various physical warehouse schemas. The participants will learn numerous types of advanced metrics, use of custom groups & consolidations, different types of filters & prompts, including relationship filters, joint element lists, dynamic dates, custom drill maps, advanced report data options, report developer tools & preferences and report Creation

Check out **Cognizant calendar** for the training details for the month

Intranet Users: <http://calendar>
 Internet Users: <http://calendar.cognizant.com>

MicroStrategy
 Teradata
 Informatica

Internet

Figure A 1.1 - Home Page of the Portal

DATA WAREHOUSING
KNOWLEDGE REPOSITORY

Home
Documents
Certification
eTrainer
NewsLetter
Upload
Tech Tips
Training Schedule
Manage Documents
Ask Us
My Subscriptions
Feedback/HelpDesk

Your Favorite Links :

MicroStrategy
Best in Business Intelligence

Teradata
a division of NCR

INFORMATICA

MicroStrategy and Cognizant: Best Practices for BI Consolidat

Upload

File* Browse...

File Classification*

Path to be Stored

File Description*

File Search Meta

Technology Category* MicroStrategy Teradata Informatica

Mandatory Fields

Done Internet

Figure A 1.2 - Screen to Upload Documents

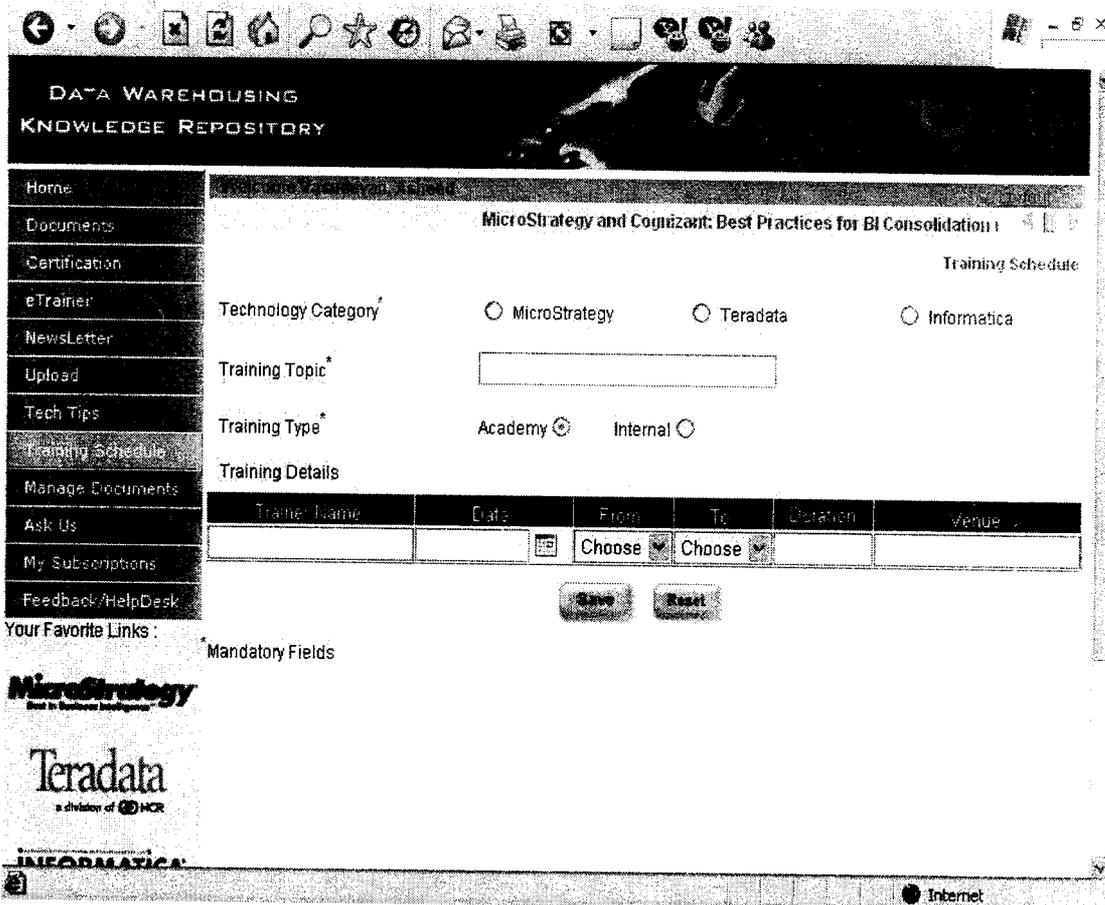


Figure A 1.3 - Screen to Upload Training Schedule

DATA WAREHOUSING
KNOWLEDGE REPOSITORY

Home
Documents
Certification
eTrainer
NewsLetter
Upload
Tech Tips
Training Schedule
Manage Documents
Ask Us
My Subscriptions
Feedback/HelpDesk

MicroStrategy and Cognizant: Best Practices for BI Consol...
Tech Tips

Tech Tip*

Technology Category*
 MicroStrategy
 Teradata
 Informatica

Effective From Date*

Effective To Date*

Save Reset

Mandatory Fields

MicroStrategy
Best to Business Intelligence

Teradata
a division of EMC

INFORMATICA

Internet

Figure A 1.4 - Screen to Upload TechTips

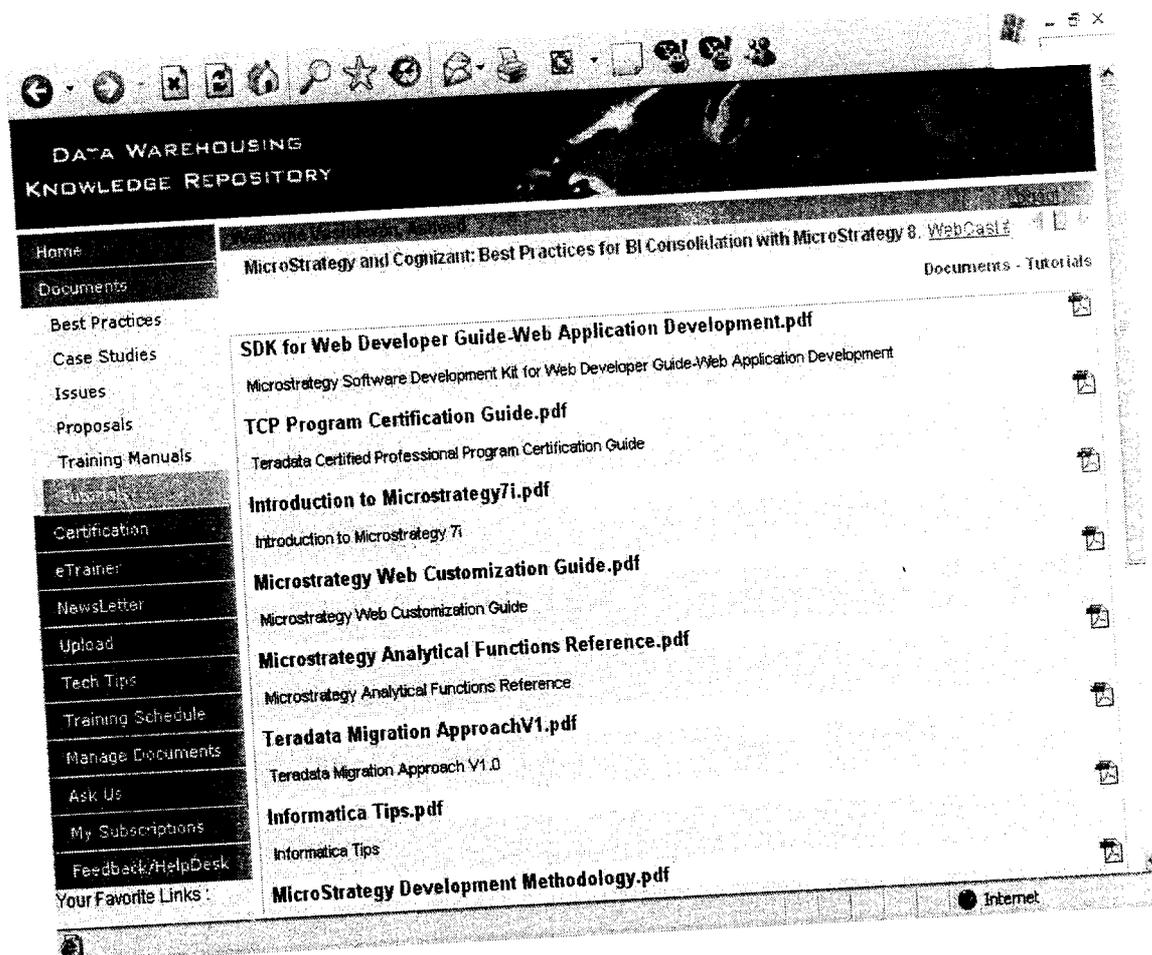


Figure A 1.5 - Screen to View Documents

DATA WAREHOUSING KNOWLEDGE REPOSITORY

MicroStrategy and Cognizant: Best Practices for BI Consolidation with F

Uploaded Documents

File Name	Technology	Classification
Informatica PowerChannel FAQ.pdf	Informatica	Tutorials <input type="checkbox"/>
Informatica FAQ		
Tuning Guidelines for MicroStrategy 7.x.pdf	MicroStrategy	Best Practices <input type="checkbox"/>
Performance Tuning Guidelines for MicroStrategy 7.x		

<< Prev 1 2 3 4 5

Change Delete Reset

MicroStrategy
Best in Business Intelligence

Teradata
a division of EMC

INFORMATICA

Done Internet

Figure A 1.6 - Screen to Manage Documents

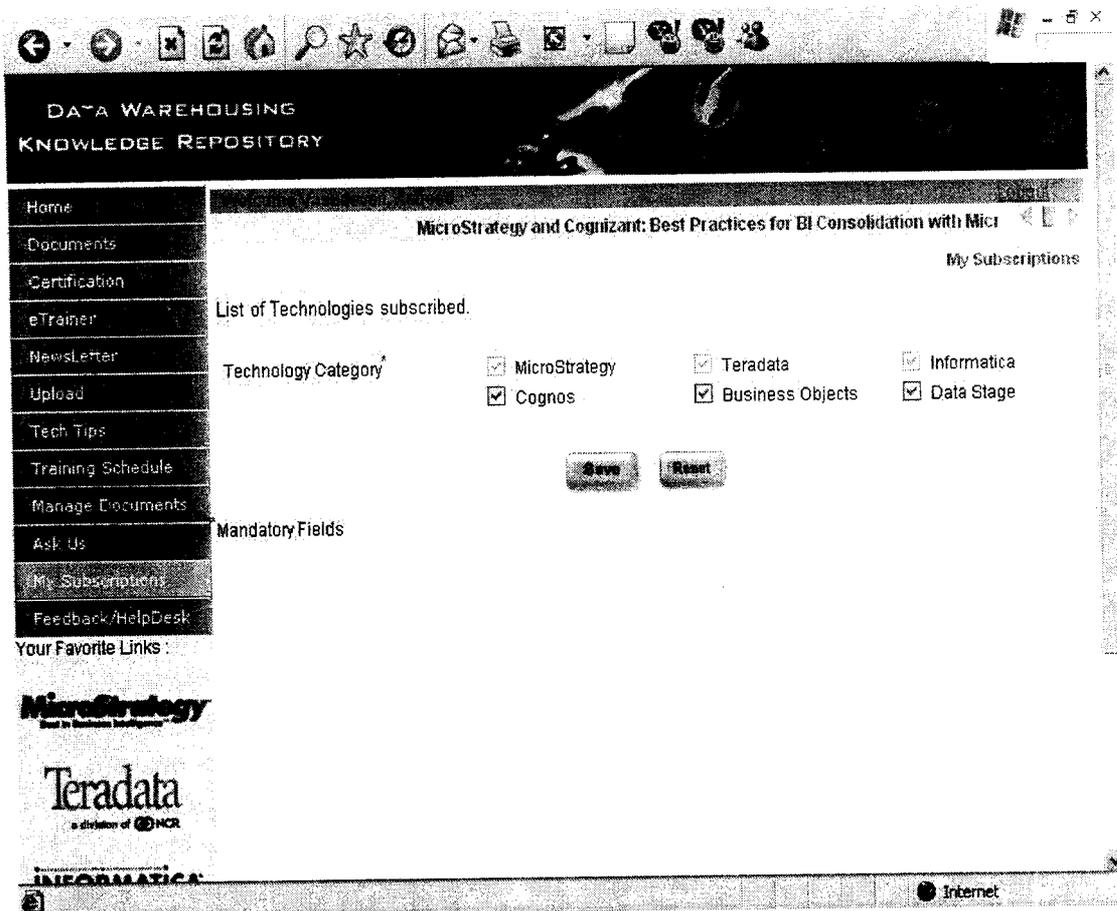


Figure A 1.7 - Screen to Edit Subscriptions

DATA WAREHOUSING
KNOWLEDGE REPOSITORY

Home
Documents
Certification
eTrainer
NewsLetter
Upload
Tech Tips
Training Schedule
Manage Documents
Ask Us
My Subscriptions
Feedback/HelpDesk

Your Favorite Links :

MicroStrategy Best in Business Intelligence
Mandatory Fields

Teradata
a division of EMCOR

INFORMATICA

MicroStrategy and Cognizant: Best Practices for BI Consolidation

Ask Us - Post Questions

Question *

Answer *

Technology Category *

MicroStrategy
 Cognos

Teradata
 Business Objects

Informatica
 Data Stage

Save Post

Done Internet

Figure A 1.8 - Screen to Post Questions

DATA WAREHOUSING
KNOWLEDGE REPOSITORY

Home
Documents
Certification
eTrainer
NewsLetter
Upload
Tech Tips
Training Schedule
Manage Documents
Ask Us
My Subscriptions
Feedback/HelpDesk
Your Favorite Links :

#1 [WebCast #2](#) Cognizant's feedback has been included in MSTR 8.0 launch press release. [Click](#)

Ask Us

- Would you like to post some questions and answers, then please click.
- For answering unanswered questions click.

1. What is Market Basket Analysis?

A technique employed in the retail industry where the contents of "typical" sets of purchases are analyzed to determine statistically which products are selling best, what kinds of affinities might exist between and among products, and similar associations.

2. What functionality is contained in MicroStrategy Administrator - Object Manager 7.x?

MicroStrategy Administrator - Object Manager allows users to perform the following actions:

- Duplicate and upgrade projects
- Copy objects within and across related projects
- Move, Delete, Rename & Search objects within a project
- Find an object's parents or children within a project

3. Are there any special requirements needed to move objects across projects?

Yes. In order to perform cross-project operations, the projects involved must originate from the same source project. In other words, the projects can only be related by the duplication of a single project. This ensures that the projects have a similar set of schema and application objects, and that the object ID's in the two projects are the same. MicroStrategy Object Manager uses the object and version ID's across the projects to perform comparisons.

Internet

Figure A 1.9 - Screen to View AskUs

2 Online MockExams Phase

DATA WAREHOUSING
KNOWLEDGE REPOSITORY

Home
Documents
Configuration
eTrainer
NewsLetter
Upload
Tech Tips
Training Schedule
Manage Documents
Ask Us
My Subscriptions
Feedback/HelpDesk

Your Favorite Links:
MicroStrategy
Teradata
INFORMATICA

Internet

ces for BI Consolidation with MicroStrategy 8. WebCast # 1 WebCast # 2 Cognizant 's feedback ha

Question Upload Page

Technology Category MicroStrategy Teradata Informatica

Category

Question

Answer Type Single Multiple User Entry

Options

1. 2. 3.
4. 5.

Mandatory Fields

Figure A 1.10 - Screen to Compile Online Test

Online Mock Exam

Introduction

- Select a Technology
- Select Category under which the questions are to be compiled on the fly
- Specify the maximum number of questions

Specify your Choice

Technology

MicroStrategy Cognos Teradata Business Objects Informatica Data Stage

Category

Available Category

MicroStrategy Reporting Essentials

Selected Category

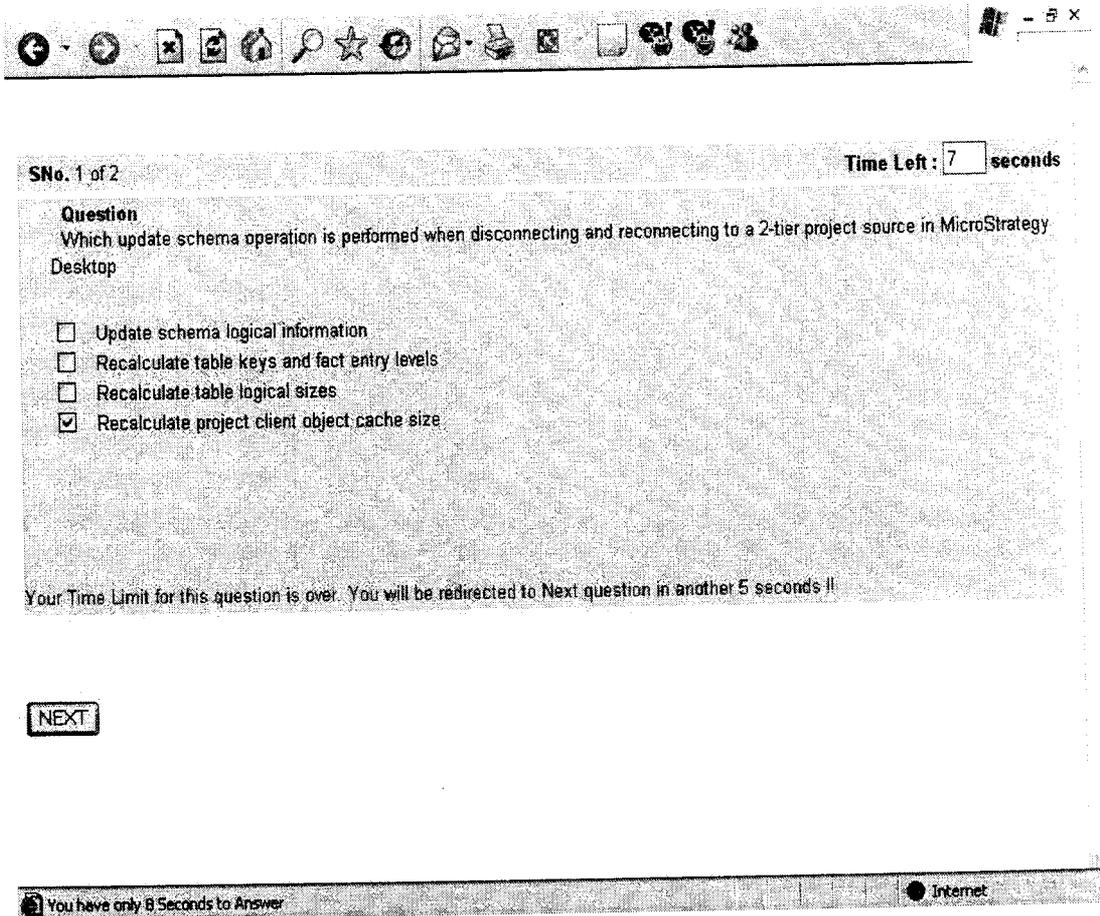
None

Enter the Maximum Number of Questions you wish to attempt

To get back to the Knowledge Repository, please click [here](#)

javascript:formSubmit() Internet

Figure A 1.11 - Screen to select exam category



The screenshot shows a web-based exam interface. At the top, there is a toolbar with various icons for navigation and editing. Below the toolbar, the question is displayed. The question number is 'SNo. 1 of 2' and the time left is '7 seconds'. The question text asks which update schema operation is performed when disconnecting and reconnecting to a 2-tier project source in MicroStrategy Desktop. There are four multiple-choice options, with the last one, 'Recalculate project client object cache size', being selected. A warning message at the bottom of the question area states: 'Your Time Limit for this question is over. You will be redirected to Next question in another 5 seconds !!'. Below the question area, there is a 'NEXT' button. At the very bottom of the page, there is a status bar with a clock icon and the text 'You have only 8 Seconds to Answer' on the left, and an 'Internet' icon on the right.

SNo. 1 of 2 Time Left : 7 seconds

Question
Which update schema operation is performed when disconnecting and reconnecting to a 2-tier project source in MicroStrategy Desktop

- Update schema logical information
- Recalculate table keys and fact entry levels
- Recalculate table logical sizes
- Recalculate project client object cache size

Your Time Limit for this question is over. You will be redirected to Next question in another 5 seconds !!

NEXT

You have only 8 Seconds to Answer Internet

Figure A 1.12 - Screen of Exam page

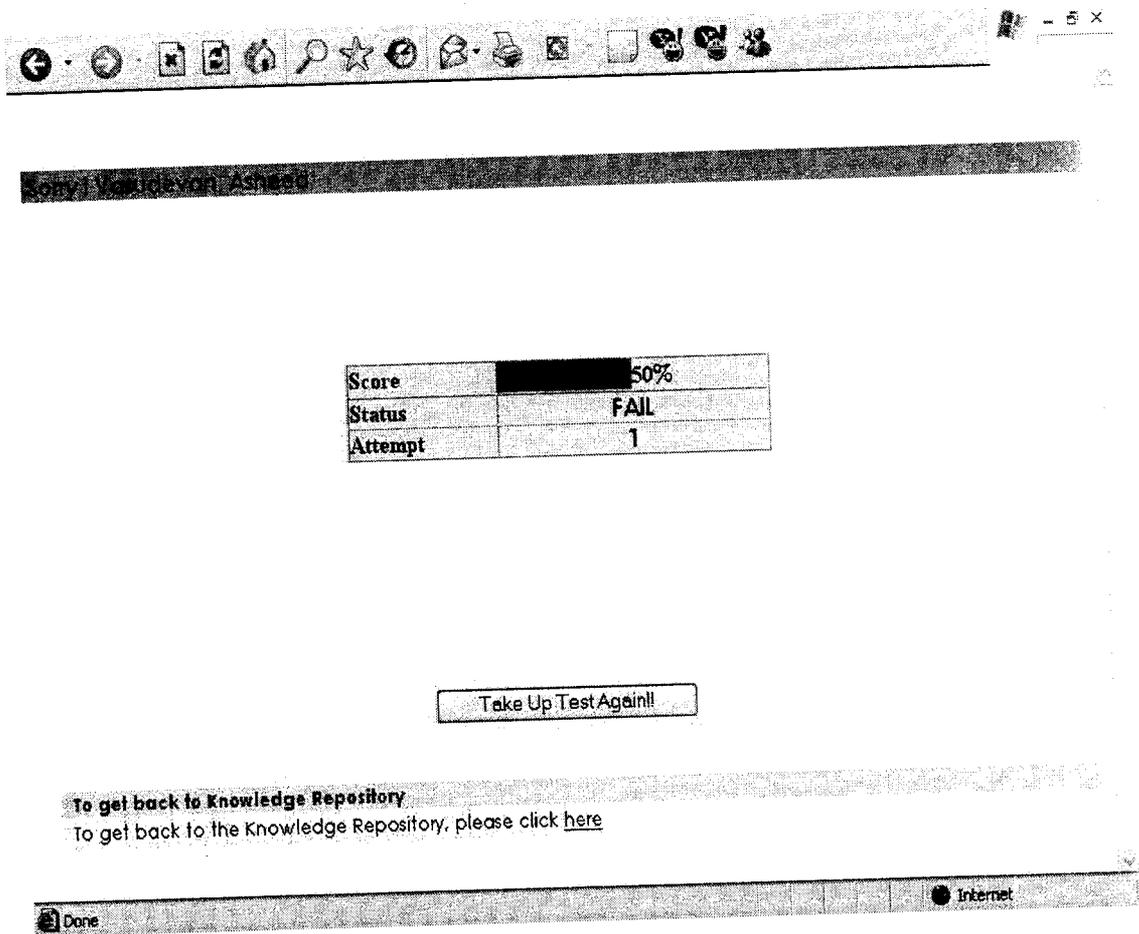


Figure A 1.13 - Screen to display result

REFERENCES

- [1] Keyton Weissinger.A, "ASP in a Nutshell, 2nd Edition"
- [2] Daniel Read, "VBScript Programmer's Reference - Second Edition"
- [3] Margaret H. Dunham, "Data Mining – Introductory and Advanced Topics"
- [4] Thomas Powell, Fritz Schneider, "JavaScript Complete Reference"
- [5] ASP Online tutorial
<http://www.w3schools.com>
- [6] CSS and XML Online Tutorial
<http://www.w3.org/>
- [7] Performance Tuning and Optimization Articles
<http://www.sql-server-performance.com>
- [8] Software Testing and Analysis papers
<http://qualityassurance.knowledgestorm.com>
- [9] VBScript Online tutorial
<http://www.gh-gold.co.uk/vbscript.php>