

CONTENT MANAGEMENT OF WEB ENABLED EXPERTS DATABASE IN THE FIELD OF SCIENCE AND TECHNOLOGY

by

Dube Sonia *
Prem Chand**
Khushwah Shivpal Singh ***
Oza Niyati ****

ABSTRACT

With the explosion in popularity of the Internet, today's computer users have become accustomed to having massive amounts of online information. The obvious consequence is the need for better Content Management System for faster and precise access to databases. Efficient content management involves content metadata, user profiling, efficient database search, data validation and web enabled access. The functionality, interoperability and application tools for content management of experts database in the field of science and technology are built herewith.

Key words: Content Management, Experts Database, Metadata, Database Search, Database Security, Servlets.

* Computer Programmer, INFLIBNET Centre, Ahmedabad

** Scientist – C, INFLIBNET Centre, Ahmedabad

*** STO-I, INFLIBNET Centre, Ahmedabad

**** Research Associate, INFLIBNET Centre, Ahmedabad

0 Introduction

The experts database system provides comprehensive information about the background, skills, and accomplishments of the individuals in the field of science and technology. This database helps in leveraging existing Research and Development efforts and conveniently provides information about Experts to peers, prospective collaborators and funding agencies.

Content Management of high volume database requires intelligent information retrieval techniques, tools and algorithms for maintaining consistency and usability of the database as presented to the users. The content management techniques adopted for building up experts database use the principles of content metadata, user profiling, controlled access and rich search techniques.

Building of Content Management System for Experts Database involves:

- ?? Evaluating the sources of contents
- ?? Creation, Integration and organization of the contents.
- ?? Identification of technology tools.
- ?? Database Organization.
- ?? Linkage of Content sources and users to create a flow of knowledge from information providers to information users.

1 Aims and objectives

The basic objectives of the content management system for Experts database are:

- ?? Quickly and conveniently provide information about experts to peers, prospective collaborators, and funding agencies in the country
- ?? Establish Communication directly with the experts who possess the expertise needed by users.
- ?? To provide current and future estimates of available resources in Science and Technology in the country.
- ?? Keep track of internal research capabilities.
- ?? Publish research capabilities externally.
- ?? Identify peer reviewers for articles and proposals.
- ?? Discover prospective collaborators and funding agencies for ongoing research projects.
- ?? Leverage existing R&D efforts
- ?? To promote and support research into areas relating Science and Technology.
- ?? To promote research of individuals and institutions.
- ?? Stay informed about the constantly changing world of basic science and scholarly research.
- ?? To create information exchanges and networking opportunities for scientists

2 Implementation

The implementation is realized using 3-tier database architecture as shown in Fig-1. Here database server is MS SQL, Apache as Web Server and Tomcat as Servlet Container. Servlets are used for performance improvement for database access as compared to process overheads incurred in case of CGI [1]. The database access and modification can be done using a web based client where the user is provided with a facility to register and create its own entry and subsequently modify his record. Bulk database can be entered and/or modified using a front end developed in Visual Basic, which provides the capabilities of registering, modifying the record, search based on specific criteria such as name, organization, specialization etc., status of database (like no. of experts entered, no. of organization entered etc.), report generation, import/export database (from SQL to CDS/ISIS and vice versa), update master records and validation of the online entered records.

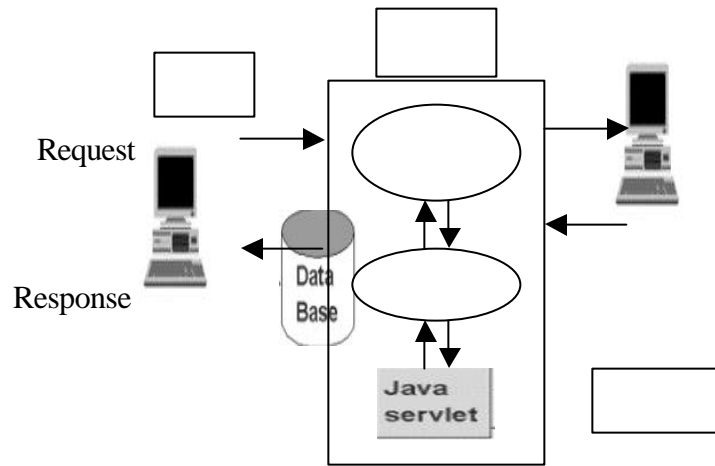


Fig-1

3 Content Management

Management of experts database includes collection of inter related data, set of programs for database access, search mechanism, content security and data validation. Content management involves both the definition of structure, which forms the core of the database, storage policy and provision of mechanisms for content manipulation.

3.1 Content Access

The content management system supports 3 different level of users with varying privileges.

- ?? Database Administrator
- ?? Authenticated User
- ?? Guest User

The Database administration job involves keeping track of valid records, sending of user-id and password using automatic e-mail mode to experts for password retrieval, record deletion and duplicate record detection to maintain the uniqueness of records. Authenticated users have privilege of modification of their records. Guest Users have only search level of privilege.

3.2 Content Search

Efficient search interface is provided to search data from database. The search techniques, as shown in Fig-2 supported by the content management system, are:

- ?? Parameter based search – In parameter based search user can search based on any given parameter or combination of parameters which forms fields in the Database.

?? Text based search – In text based search user is asked to give a text and then this text is searched in the metadata for retrieval of matching entries.

Parameters supported for parameter based search are:

- ?? Expert Name
- ?? Organization Name
- ?? Subject Specialization
- ?? City
- ?? State
- ?? Geographical Region
- ?? Language Knowledge

Complex search can be built by using two or more Boolean operators (AND, OR, NOT) in parametric search. Use of Boolean operator helps in either narrowing or expanding the search.

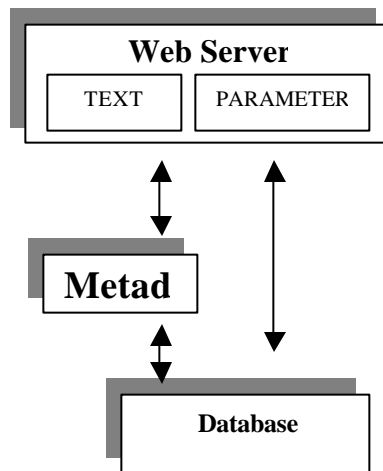


Fig-2

Metadata helps the user to specify additional document attributes, which describe the essential aspects of the text in the database [2]. Metadata is designed to improve the search performance of text based search methods for large volume of databases. Managing the content metadata involves extracting the metadata information from the text submitted by the expert, and then storing this information, which can be used during the search. Metadata can be extracted either manually or can be done by several automatic methods available. Metadata can either be stored along with the expert's records in the same database or can be stored in a separate database with a link to the relevant record. Metadata can also be used as a parameter to judge the quality of the expert's record.

3.3 Content Security

The Content Security is implemented at record level by allowing only authenticated user to change/update his record in the database [3]. Each Record is validated against known

rules before it is updated/entered in the database. Apart from the record level access protection, following security levels are implemented:

Application Security: Experts database can be accessed either using a web-based client or VBClient. The application security for web based approach includes controlled access to database. Based on user authentication, different levels of privileges are granted to the user for access to specific web pages. The VBClient is designed for Bulk data transfer, general database administration operations and requires authentication to gain administrative privileges.

Data Security: Data Security for Experts database is implemented at the Database server level, which is MS SQL. Microsoft's SQL Server is a client/server database engine. The server side of the application provides security for database and requires clients to authenticate for access to the data.

Content security ensures that only authorized user is actually allowed to access the content database.

3.4 Stored Procedures

Use of Stored procedure in a content management system helps in improving the speed of access, security and data independence [4]. A stored procedure consists of one or more SQL commands encapsulated in a single module. Stored procedure provides speedy access to database as compared to standard SQL queries, which are interpreted and parsed at runtime. Use of stored procedure significantly improves the security aspects of the database as it is no longer necessary to allow any user to have modify/delete access rights on the database. Providing execute permissions of these procedures serves the purpose. Stored procedure makes the code database independent and hence more portable as use of stored procedures eliminates the need of embedding the SQL queries within the program.

4 Online Registration and Modification

Individual experts can register and modify their profiles, online. Online registration requires user to provide his preferred UserId and password, which can be used by him at a later time to modify or update his profile. Registration of User is validated against Name, Organization and City in order to avoid duplicate entry creation in the database. To modify profile, expert has to authenticate himself by providing valid UserId and password and once authenticated, he has privileges for modification and Updation of his profile. The database includes a field for recording the IP address of the machine, which was used to modify or update the database. This helps in keeping track of experts origin. The database also contains a field for recording the date when the database was last modified or updated. This information is used to send reminders to the expert for record updation.

5 Content Validation and Quality

Content Validation is an essential component in content management system as it ensures the correctness and validity of the information being provided to the users. Content Validation includes validation of essential fields to be non-null, automatic detection of duplicate records and numeric field validation for number formats. Content validation checks the types of text and fields when a form is submitted. In order to ensure and maintain the database quality, all records submitted online are manually verified, before their profiles are published on the web. This prevents unauthorized users to include their profiles in the database. Concept of Information ageing [5] is used to ensure removal of old and outdated records from the database.

6 Conclusion and future Scope

As in any content management system, the success and level of utilization of the database is highly dependent on the quality, quantity and precise access to the database and hence initiatives have been taken to create profiles of experts in various R&D organizations. New Search techniques are to be incorporated for speedy access to the database. Web enabled access to users have been provided over Internet to extend the availability of this information both for content creation and retrieval. In order to make this web-site more interactive, discussion fora and chat rooms are planned, where the users can post their queries or can online chat with the experts.

7 Acknowledgement

Experts database in Science and Technology (EDST) is a collective effort of INFLIBNET and NISSAT. It is a NISSAT funded project under the title “Developing and Managing a web enabled Database of Experts Manpower in the field of Science and Technology”. Under the project, INFLIBNET has mandate to collect, update, maintain and host the profiles of scientists and other senior faculties working in leading Research and Development Labs, universities, Science and Technology departments of Central and State governments in the country.

The Authors thankfully acknowledge the encouragement and support received from Dr. T A V Murthy, Director, INFIBNET Centre, Ahmedabad and Shri S. M. Salgar, Scientist-G, INFIBNET Centre for his constant interaction and assessing the progress of the work.

8 Bibliography

- [1] MOSS (KARL). Java Servlets. 2nd ed. 1999. Tata McGraw-Hill Edition. p 2-3.
- [2] Principles of Intelligent Content Management
<http://www.intelligententerprise.com>
- [3] KORTH (HENRY F) and SILBERSCHATZ (ABRAHAM). Database System Concepts. 2nd ed. 1991. MaGrw Hill Edition. p 525-533.
- [4] Tenrox Enterprise Software Security Features
<http://www.officepurchasing.com/en/technology/security.htm>
- [5] DAVIS (GORDON B) and OLSON (MARGRETHE H). Management Information System, Conceptual foundations, structure and development. 2nd ed. McGraw Hill. p 223-225