

**CREATING HEALTH SCIENCES THESIS INSTITUTIONAL  
REPOSITORY MODEL: AN APPROACH AT AIIMS,  
NEW DELHI, INDIA.**

***Sangeeta Narang\**, *S. Kailash\*\**, *S. K. Meher\*\*\** and *Jagdish Arora\*\*\*\****

*\*Librarian, B.B. Dikshit Library, AIIMS, New Delhi, India*

*\*\*System Analyst, Computer Facility, AIIMS, New Delhi, India*

*\*\*\* Sr. Programmer, Computer Facility, AIIMS, New Delhi, India*

*\*\*\*\*Librarian, Indian Institute of Technology, New Delhi, India*

**Key words :** Medical Libraries, Thesis and Dissertation, ETD Standards,  
Institutional Repository

Address for correspondence :

**Sangeeta Narang**

Librarian

B.B. Dikshit Library

All India Institute of Medical Sciences (AIIMS)

Ansari Nagar, New Delhi-110029

India.

Ph. : 91-11-26588500, 26588700 Extn. 4595, 4371

Fax : 91-11-26588663

Email : sangitasnarang@yahoo.co.in

## ABSTRACT

**Introduction:** One of the major concerns today is to put raw information into useful knowledge. Theses and dissertations largely constitute an unrevealed/ hidden information. But with the infusion of the technology their content is reaching the global audience. It is enhancing readership, quality research, minimizing duplication of efforts and much more.

**Background:** The All India Institute of Medical Sciences (AIIMS), New Delhi, India is one of the premier institute of the country providing patient care, research and teaching. Each year B.B. Dikshit Library of the institute receives on an average 250 theses and dissertations from post graduate and doctoral students in the multispeciality areas of health sciences. Presently the theses catalogue is available to the users via online public access catalogue (OPAC) in the library. In the future it is planned to put theses content on a network to the users worldwide.

**Methodology:** The paper is based on a questionnaire to determine the utility and feasibility of the system. Based on a survey result and literature search a model is devised. It includes: ●standard mechanism and means for submission, archiving and accessing theses, ●comprehensive document retrieval, indexing and storage, ●full text and parametric search, ● access and security rights at various levels and ●unicode support. It also covers system requirements – hardware, software, operating systems, database, application server and web server.

**Conclusion:** Today, all the thesis and dissertation are created electronically. Therefore, it is essential to device such an information system model through which information flows everywhere with just the click of button. Health sciences demands real time delivery of information to the right sources at right places.

## **1. INTRODUCTION**

Theses and dissertations are raw, rich and unique source of research information, often the only source that does not find its way into various publication channels. Doctoral dissertations are manifestation of result of four to five years of intense work involving huge investments. Thesis reflects quality of research work conducted by a student over a period of time. They are characterized by the original work carried out by a student to achieve a degree.

Today, most of the available bibliographic databases do cover research articles from national and international journals, conference proceedings, but still do not index theses and dissertations except Dissertation Abstracts International.

## **2. INITIATIVES IN INDIA**

Theses collections in most of the Indian libraries are kept in closed access, making it difficult for other students to access them. It remains an unrevealed under utilized asset leading to unnecessary duplication and repetition that, in effect, is the anti theses of research and wastage of resources.

In the country UGC has laid down regulatory frame work to improve quality, accessibility and availability of Indian theses to bring up uniform standards for creating metadata. There are number of agencies that are involved in collection, compilation and presentation of bibliographic content (meta data) of theses in India. Some of the major initiatives are:

- Association of Indian Universities(AIU) publishes a list of theses submitted in various universities in its weekly publication “University News”
- The Information and Library Network (INFLIBNET) and Developing Library Network (DELNET) host databases of bibliographic records of PhD theses submitted to various universities in India. They consist of 1,40,000 and 3,953 bibliographic records respectively.
- Currently, Vidyanidhi is a nationwide effort in digitizing theses and dissertation supported by Ford Foundation and Microsoft India. It hosts more than 500 theses in full text and 85,000 bibliographic records. Vidyanidhi is a member of Network Digital Library of theses and dissertations (NDLTD), Virginia Tech (USA).
- Various other institutes namely Indian Institute of Sciences (IISc), Indian Institute of Technology (IITs) are building their digital theses collection.

Inspite of number of sporadic efforts mentioned still we do not have a comprehensive and authentic source of information.

### 3. INITIATIVE AT AIIMS, NEW DELHI

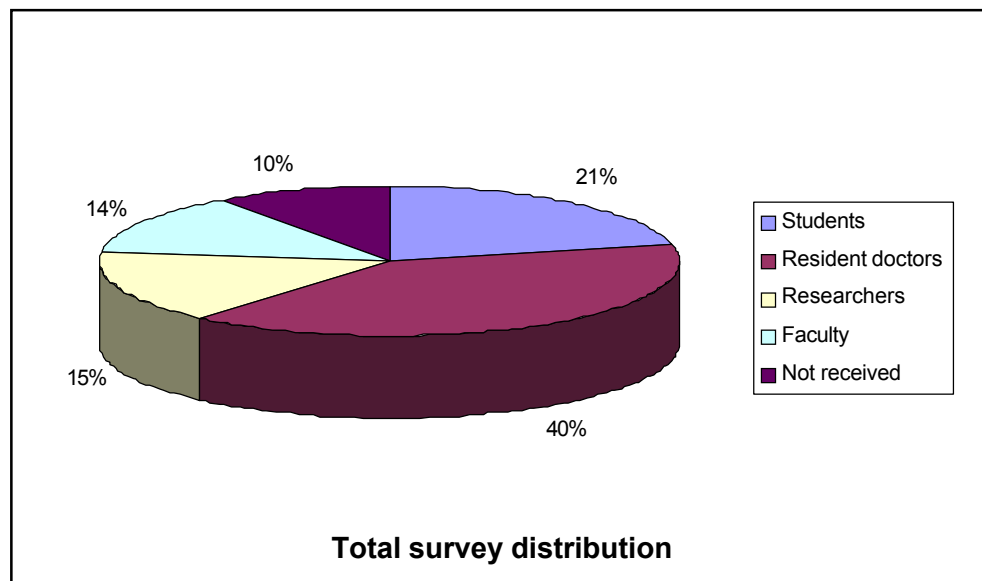
The All India Institute of Medical Sciences (AIIMS), New Delhi, India established in 1956, as an institution of national importance. It provides comprehensive facilities for teaching, research and patient care. Each year B.B. Dikshit Library of the institute receives on an average 250 theses and dissertations from postgraduate and doctoral student in the multispeciality areas of health sciences. Presently, the theses catalogue is available to the users via online public access catalogue (OPAC) in the library. It has 4307 bibliographic theses received so far.

### 4. USER'S SURVEY

In order to determine the utility and feasibility of developing theses institutional repository model questionnaire was developed. A pilot study of the survey was done within the campus. Respondents were requested to return the filled questionnaire in the library's office.

#### 4.1. Findings

A total of 140 of the 165 surveys distributed were received from 33 undergraduate students, 62 resident doctors, 24 researcher and 21 faculty members.

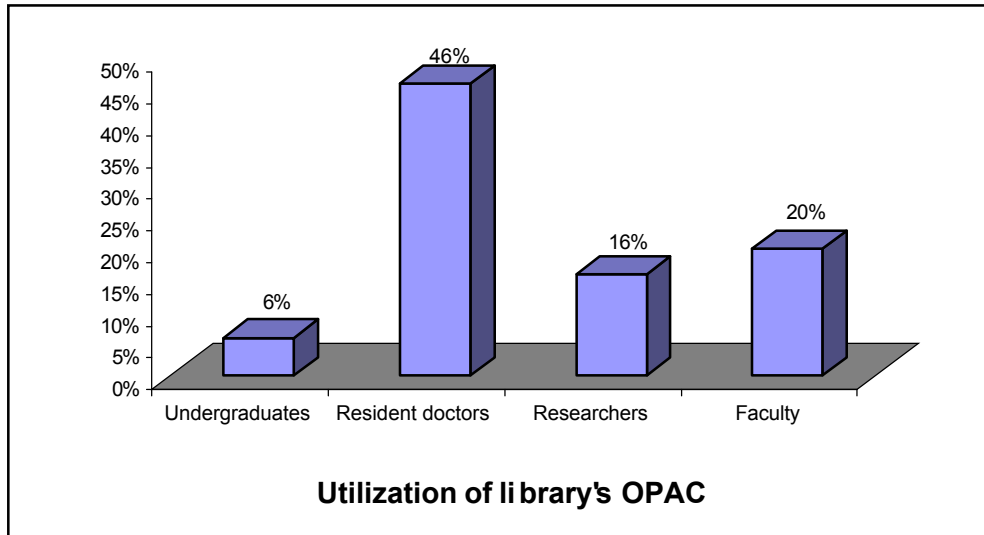


Findings revealed that electronic resources have become an inescapable source of information. Each one of them rated it as highly important.

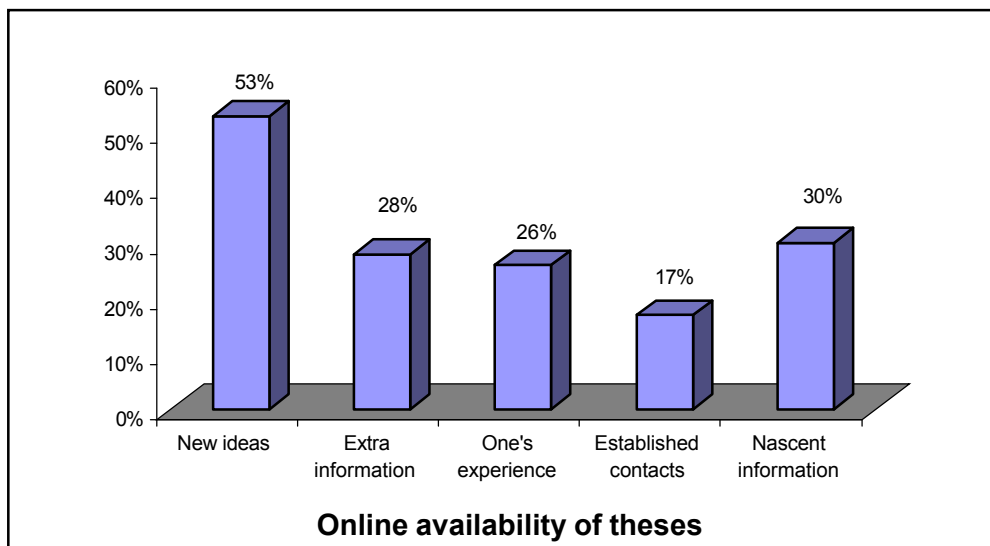
As far as theses consultation in the library is concerned it came out to be very less among the users. Only 24% responded to its active use.

Seventy four percent suggested that library should have only electronic format while rest suggested that along with electronic, print should also be acquired.

Eighty seven percent revealed that theses should be available online; the library's OPAC has not been extensively utilized. Six percent undergraduates, forty three percent resident doctors, sixteen percent researchers and twenty percent faculty members have utilized this in-house library service.



Apart from these users suggested that if theses will be available online the it will impart new ideas / directions (53%), provide information about what others are doing (28%), enable to provide one's experience (26%), establish more contacts (17%), provide information not received from other sources (30%).



## **5. RESULTS**

Findings revealed that users need electronic proliferation of theses and dissertation institutional repository. Based on it a model is devised.

## **6. THE PROPOSED MODEL CONSTITUTE**

### **6.1. Standard Mechanism and means of submission**

The preparation of electronic thesis and dissertation (ETD) involves making of the electronic copy of the thesis / dissertation. Before digitizing thesis it is required to have student's approval and consent regarding their right to archive and distribute their doctoral work. The most popular and preferred format to accept e thesis is a portable document format (PDF) or to buildup e thesis websites such that researcher can logon to ETD server- enter bibliographic information and upload text. A file name is also desired so as to have unique name for each file. Initial pages like acknowledgement, certificate clubbed into one single file as 'intro.pdf'. Rest is further divided chapter wise as text.pdf

### **6.2. Metadata Support.**

To have efficient, effective organization, access and retrieval of information contents in a digital library, it is essential to have a metadata support. It acts as a browsing and a search interface with well defined attributes of a digital object. There are various metadata schemes like machine readable catalogue, (MARC), Dublin core, ETD-MS etc.

### **6.3. Open Architecture Initiative**

Once the digital objects are organized in predefined categories search/ browsing interfaces, it is essential to have an open archive initiative compliant software to built up a repository. One such standard for metadata harvesting is open archive which is interoperable. It is a means to enhance access to e-prints/ pre-print archives. There are several OAI-compliant repositories worldwide. (<http://www.openarchives.org/>)

### **6.4. Various Access Methods**

An institution has option to put theses world wide to the entire research community or to restrict it to a few privileged users. The types of access that can be given are:- World wide access , Campus-only access, Temporary Restricted Access, Mixed Access, Withheld Access.

### **6.5. Unicode support**

Since Unicode allows the data to be transported through many systems without corruption and enables a single software product or a single website to be targeted across multiple platforms, languages and countries without re-engineering . So it is essential to have Unicode standard for thesis and dissertation.

## **7. DISCUSSION**

- The library's OPAC is available within the library; it does not find much usability as doctors have a busy schedule. But once theses and dissertations will be available online on web it will add on to their flexibility, convenience and comforts.
- Health sciences demands real time delivery of information to the right sources at right places. As it encompasses disease diagnosis, treatment, prevention and control and health promotion, therefore it is essential to have a reliable and quality assured databases like ETD.
- Institutional repository / e-print archive enables to acquire, preserve and provide institutional intellectual asset to wider community, thereby enhancing and encouraging research.

## **8. MINIMUM SYSTEM REQUIREMENTS**

### **Server**

#### HARDWARE

Server Grade Machine with minimum 512MB RAM , 80GB HDD

#### SOFTWARE

OS: Windows: NT server, 2000 Server / Sun Solaris/Redhat Linux.

Database: Oracle, SQL, PostgreSQL or MySql.

Application Server: BEA Weblogic, Oracle 9iAS.

Web Server : Apache

### **Client**

#### HARDWARE

Pentium, with minimum 64 MB RAM, 40 GB HDD

#### SOFTWARE

OS: Windows 2000, XP, NT

Other Software: MS Office 2000, Internet Explorer, Netscape

## 9. TECHNOLOGICAL OPTIONS

A number of software solutions are currently available for building digital publishing systems that can support different document types including electronic thesis and dissertation. Several software are open source, OAI-complaint solutions that allow its users to build an open-standards compliant interoperable ETD repository. Few of them are:

- **ETD-Db:** The ETD database (ETD-db) developed at Virginia Tech. (<http://scholar.lib.vt.edu/TED-db>)
- **DSpace:** Developed in partnership between Hewlett-Packard(HP) & Massachusetts Institute of Technology (MIT). (<http://dspace.org/>)
- **Theses Alive Plug-in for Institutional Repositories (TAPIR):** Developed at Edinburgh University Library(EUL). (<http://www.thesesalive.ac.uk>)
- **E-Prints:** It is an open source digital library software package developed in the University of Southampton. (<http://software.eprints.org/>)
- **Flexible Extensible Digital Object Repository Architecture (FEDORA):** It is an open source digital object repository system using public APIs exposed as web services. (<http://www.fedora.info>)

## 10 MAJOR ISSUES

1. **Plagiarism:** A risk of Plagiarism is one of the important concern that most of the students and faculty faces. The availability of documents in electronic format makes the work easier for authors to copy.
2. **Intellectual Property Rights (IPR) and Copyright.:** The owner of copyright of a book or any other written document belongs to its creator or author, irrespective of media used for its presentations, i.e. paper or electronic. The author of an electronic thesis or dissertation is its copyright holder and thus owns the intellectual property contained in it.
3. **Obsolete Technology:** In the archival prospective it is feared that if the software becomes obsolete and goes out of use there is no way one can retrieve the document.

## 11. CONCLUSION

The system so proposed will meet the requirement of our institute in particular and with improvement in the design will make the institutional repository retrieval much easier. Once Electronic Theses and dissertation (ETD) are available onto the users it will boost e-publishing technology which may include links, hyperlinks in the body of document and will save time and labour.



## 12. REFERENCES

1. University Grants Commission, New Delhi,2005: Regulations Framework for Creation of Metadata for Doctorial Thesis and Promoting Electronic Submission of Thesis: Guidelines and Basic Tutorials.
2. Rajashekar T.B. (2004): Establishing Institutional Repositories (IR) in Indian Universities. Position paper (draft) prepared for the UGCETD framework committee.
3. <http://www.vidyanidhi.org.in>
4. <http://www.unicode.org/standard/>

To : Medical Students and Health Professionals  
From : BBD Library, AIIMS, New Delhi

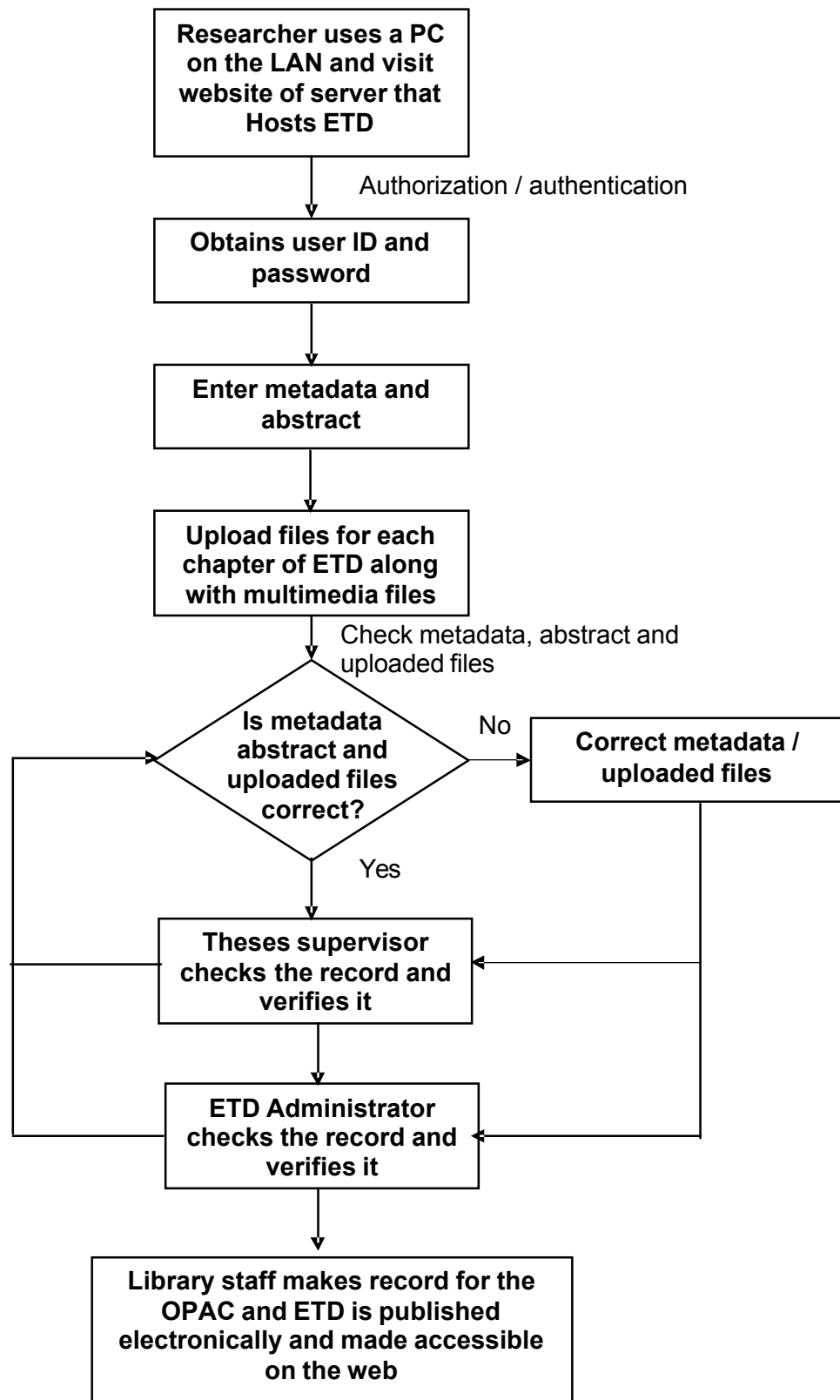
Information Technology is involved in every sphere of human life. In order to create **“Health Sciences Theses Institutional Repository”**, I wish to conduct a survey. I would appreciate if you please take few moments to help me in conducting this research. Looking forward for your Co-operation.

### Questionnaire

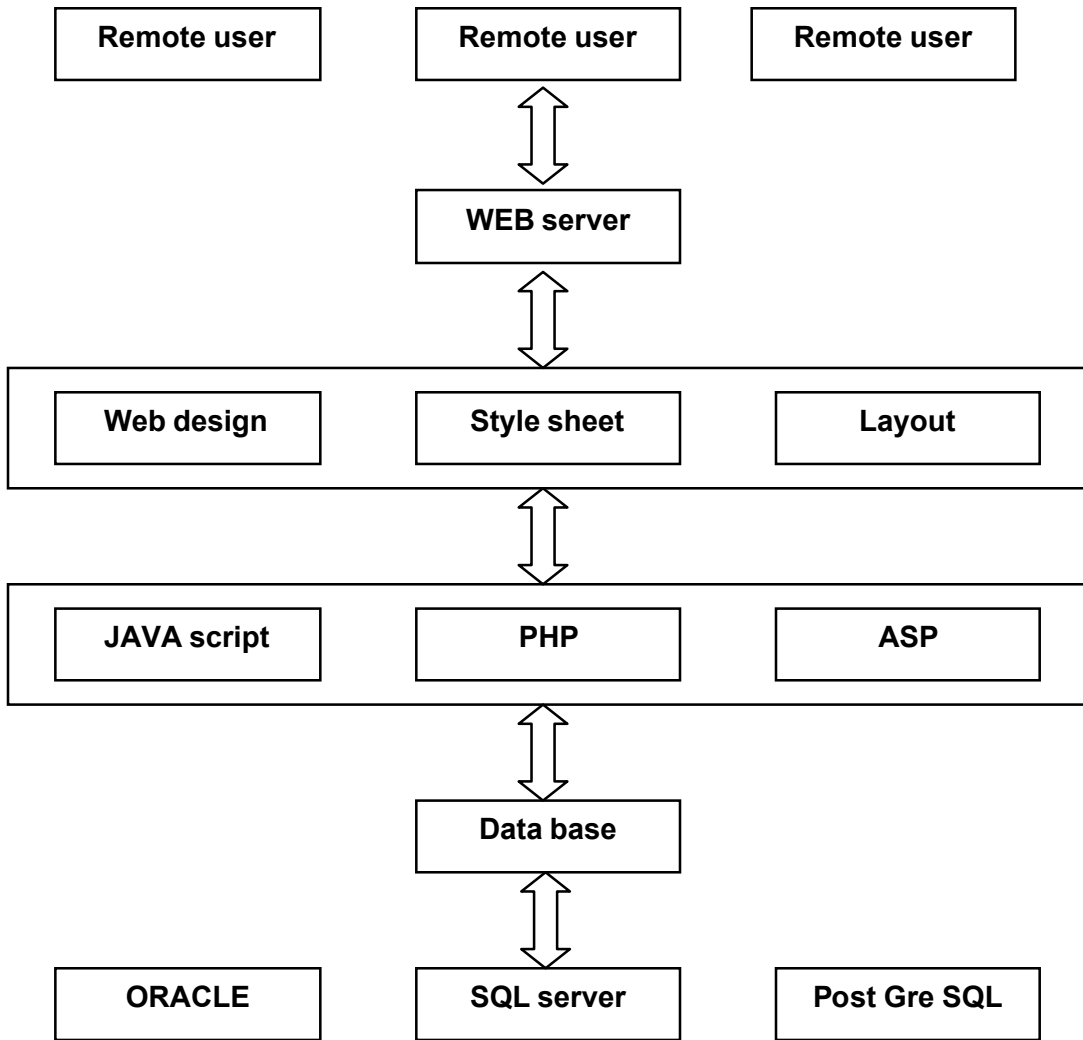
1. Your designation  
(a) Student  
(b) Resident  
(c) Faculty  
(d) Researcher  
(e) Other  
(specify)\_\_\_\_\_
2. Your Qualification  
(a) MBBS  
(b) MD/MS  
(c) DM  
(d) Ph.D  
(e) Other  
(specify)\_\_\_\_\_
3. Your e-mail address \_\_\_\_\_
4. (a) Your department \_\_\_\_\_ (b) Specialization \_\_\_\_\_
5. How do you rate the importance of the electronic resources to your work?  
(a) High importance  
(b) Somewhat importance  
(c) Little importance  
(d) No importance  
(e) Don't know
6. Do you consult theses in the library to support your work? -Yes / No
7. How often do you refer thesis in a month ?  
(a) Once  
(b) Frequently  
(c) Rarely  
(d) Never
8. What theses format should library accept?  
(a) Print only  
(b) Print & Electronic  
(c) Only Electronic  
(d) None
9. What is your opinion regarding theses collection in the library  
(a) Adequate  
(b) Inadequate  
(c) To be improved  
(d) Uncertain

10. Do you think library should digitize existing bound theses? - Yes / No
11. Once available in electronic format, do you suggest it should be available on line? - Yes / No
12. Do you think online theses will be useful to the readers? - Yes/No
13. Have you ever used Library's Open Public Access Catalogue (OPAC) for theses retrieval?  
 (a) Frequently  
 (b) Often  
 (c) Rarely  
 (d) Never
14. What do you suggest online theses should have ?  
 (a) Bibliographic content only  
 (b) Abstract  
 (c) Full text  
 (d) None
15. How do you prefer to read/make correction in theses? - On computer / On print
16. Which other database do you use for your research work ?  
 (a) Dissertation Abstract International  
 (b) Universities' web sites  
 (c) University news  
 (d) Others \_\_\_\_\_
17. Do you think online theses will help to  
 (a) Provide new ideas / direction  
 (b) Provide information about what other are doing  
 (c) Enable to provide one's experience  
 (d) Establish more contacts  
 (e) Provide information not received from other sources  
 (f) Any other (please specify) \_\_\_\_\_
18. Are you able to keep with advances in your field? - Yes / No
19. What are the different problems faced by you in obtaining information and keeping up with advances in the field?  
 (a) Lack of time  
 (b) Information scattered in too many sources  
 (c) Information is too vast  
 (d) Don't have access to library  
 (e) Library has inadequate technology  
 (f) Any other \_\_\_\_\_
20. Do you have additional comments or recommendations?  
 (a) \_\_\_\_\_  
 (b) \_\_\_\_\_  
 (c) \_\_\_\_\_

***Thank you for your participation***



**ELECTRONIC SUBMISSION OF THESES AND DISSERTATION**



**DESIGN DIAGRAM**