CHAPTER 3

RESEARCH METHODOLOGY

- Introduction
- Selection of Sample
- Data Collection Tools and Methods
- Questionnaire Design
- Data Collection Procedure
- Data Analysis Techniques
CHAPTER 3

RESEARCH METHODOLOGY

3.1. INTRODUCTION

This chapter narrates the system of models, procedures and techniques that have been adopted and followed to systematically study the result of the research problem “Digital Library Initiatives in South India: Prospects and Issues”. It is a step-by-step illustration of the systematic procedures by which the researcher starts from the initial identification of problem to its final conclusions. The objective has been to carry on the research work in a scientific and valid manner. The method of research provides the tools and techniques by which the problem is approached.

The present study is focused on the prospects and the multitude of issues and challenges being faced by the ongoing Digital Library Initiatives in South India. Hence, this study systematically assesses the various related components and issues covered by the topic of the study. The study is based on the data collected on 45 digital library initiatives with special reference to South India. The digital libraries developed in various types of organizations including educational and research institutions, are also included in the study. A sample survey of such libraries have been conducted. In order to achieve the objectives of the study, data collection methods included survey method using questionnaire, public domain resources, email/phone correspondences, observation and personal interviews. A near-comprehensive and an exhaustive literature survey have been carried out on the topic of research and other related fields. For this purpose LISA database, and other reference sources like bibliographies, online
information resources, conference proceedings, library science journals, books etc. were scanned with a view to collect relevant literature.

3.2. SELECTION OF SAMPLE

The sample of the research study consists of significant ongoing digital library initiatives in South India, which are reported in literature and publicized to the outside world. Out of a total of 41 institutions, 45 digital library initiatives were identified for the study. For the present study the term ‘Digital Library Initiatives’ include focused Digital Collections, Institutional Repositories and Electronic Theses and Dissertations (ETDs), developed using Digital Library and IR software packages. It may be noted that three institutions out of 41, viz., the Indian Institute of Management Kozhikode, the Indian Institute of Science Bangalore and the Indian Institute of Technology Madras have more than one digital library initiatives.

3.3. DATA COLLECTION TOOLS AND METHODS

The data sources for this study constitute the organizations / institutions / projects / digital library initiatives in South India who are involved in the development and / or upkeep of digital libraries. Data regarding status of digital library initiatives and various issues involved in the design and development of these digital library initiatives have been collected by using the following methods:

(i) A questionnaire (Appendix-II) targeting select digital library initiatives in South India;

(ii) The websites of the libraries forming part of universe of the study;

(iii) E-Mails and attachments;

(iv) Visits to select digital libraries in the region for personal observation;
(v) Interview with the heads/ librarians of these initiatives;

(vi) Other documentary sources on the library’s details, viz., brochures, flyers, guides etc. which could support the datasets, being considered for the study.

3.3.1 Questionnaire Design

A well planned, elaborate and structured questionnaire has been administered for data collection of the research study. The questionnaire has been conceived as a systematic compilation of questions logically related to problem under study. Bearing the research objectives in mind and equipped with an extensive literature review, meetings and discussions with the library community were conducted prior to the conceptualization of the possible queries / items in the questionnaire. The main purpose of this questionnaire has been to collect a detailed and elaborate account relating to the design and development of digital library initiatives in South India.

The data collected have been categorized into the following twelve broad significant areas of digital library development aspects:

A. General;
B. Digital Library Initiatives – General;
C. Information Technology Infrastructure;
D. Collection Development;
E. Human Resources for Digital Library Initiatives;
F. Accessibility;
G. Metadata and Interoperability;
H. Preservation and Archival;
I. Copyright;
J. Digital Library Software;
K Funding

L. Challenges Involved in Digital Library Development

Section A in the questionnaire was quite general. Seven questions asked in this section, aimed to get the name, year of establishment, type of institution / management (parent organisation), field of discipline of the institution/organisation, and the name and address and institution’s website.

Section B was labeled digital library initiatives – General, had eleven questions asking the details of the digital library initiatives, such as name of the digital library initiatives, year of commencement of the initiatives, URL, name of the head/person in charge of the digital library, involvement in the creation of digital library, type of digital library initiatives, factors which prompt and motivate in the creation of the digital library initiatives, project mission/ goal/objective, future plans with respective of strengthening, expansion and extension etc., potential benefits of the digital library initiatives.

The questions in Section C were meant to collect data on information technology infrastructure. A total of 10 questions were presented in this section, which includes, different types of servers, Availability of LAN, type of communication model of data network, nature of connectivity, name of service provider, bandwidth of internet connectivity, problems faced with the use of campus network, availability of workstations., status on the connectivity to National Knowledge Network (NKN), availability of ICT strategy and policy.

The questions in Section D were intended to collect the details of the collection development. A total of 15 questions were asked on total collection, type of collection, sources of digital documents, media types, whether the collection is static/ dynamic,
type of digital material, nature of acquisition, nature of collection, formats, digitization, volume and magnitude of the collection, mass digitization personnel attending the work, challenges involved in digitization, availability of collection development policies.

**Section E** was meant to collect data regarding human resources for the digital library. 10 questions were asked in this section on strength of library staff, number of personnel directly involved with digital library, who is/are handling the digital library at the implementation level, does the institution have a digital library implementation committee, who is the member of the committee, personnel with computer skills, does the digital library offer IT training courses to the staff, is the IT training outsourced and done by outside agencies, are there any system of specific courses/training given. Training opportunities or programs on the creation or use of digital libraries, challenges in terms of technical knowledge capabilities in digital library creation/collection development operation.

**Section F** was framed so as to get the details on the accessibility. There were four questions asked on how do you host your collection, how do you provide access to your digital collection, how do you provide access control to your digital library, what are the information organization model followed by the digital library.

**Section G** was meant for assessing the metadata and interoperability standards used in creating the digital libraries. Three questions were asked on the type of metadata, metadata schemas and interoperability standards.

**Section H** was about to collect the details on the preservation and archival. A total of eight questions were asked on their responsibility for preserving the digital
materials, presence of formal policy, nature of preservation, preservation methods, type of digital storage media used.

Section I dealt with copyright and there were three questions were asked on various aspects of copyright. They include copyright permission, measures applied to manage the copyright/digital rights of e-resources and rating of copyright.

Section J sought details on digital library software and 10 questions were asked about different aspects of digital library software such as type of software used, name of open source software, reason for opting open source software, plan to migrate to another software, if yes the reason, satisfaction with the features of the software, rating of the features of the software to other software, implementation of any modifications/customization for the software chosen and major obstacles in successfully adopting OSS.

Section K meant to collect information regarding budget and sources of funding for the digital library initiatives. There are total three questions were asked about nature of funding, separate budget allotted for the digital library, total cost involved in setting up of the digital library.

Section L was labeled ‘Challenges Involved in Digital Library Development’ and listed thirteen possible related challenges that a digital library could be facing in creation and developmental such as lack of funding, lack of staff skills; relevant expertise, lack of technical knowledge, lack of training, inadequate infrastructure, lack of management support, security problems, identification of valuable and reliable resources, copyright/digital rights management, attitude of publishers, lack of standards, lack of content, and lack of policies. Respondents are asked to mark the challenges depending upon their problem by using five point Likert scale.
3.4. DATA COLLECTION PROCEDURE

Questionnaires were sent to 55 select libraries with a view to obtain information on various aspects of the digital libraries. Out of the total questionnaire distributed 45 digital libraries (81.11%) have responded. The questionnaires were sent by post along with a self-addressed and self-stamped envelope. Reminders were sent to the non respondents after a gap of two months. E-Mail reminders were sent repeatedly even afterwards. Web sites / home pages of the libraries under study were visited consistently with a view to gather / update the relevant data. In addition to the above tools of data collection, visits, personal interactions, discussions, e-mail correspondences, interview, observation, publicity materials and websites were adopted to supplement the collection process towards requisite data for the research.

3.5. DATA ANALYSIS TECHNIQUES

The collected data were processed by using MS Excel and SPSS Software. Quantitative study and percentage analysis along with Chi-square test, weighted average score and ANOVA wherever necessary, were employed to analyze the obtained data. The analysis was presented and interpreted by using Tables, Charts and Figures.

3.5.1. Chi-square Test

Chi-square symbolically written as X2 is a common test for analyzing data from surveys. It is an inferential statistical test that is used to examine relationship between two variables with nominal or ordinal data. The Chi-square value measures the discrepancy between the observed frequencies and the expected frequencies. The larger
the Chi-square score, the larger the discrepancy, and the more likely that the two variables being studied are related. If the calculated value of Chi-square is less than the table value it indicates that the difference between actual and observed frequencies is due to chance of variation and can be ignored. The Chi-square calculation can be summarized into a mathematical formula, which can be applied to contingency tables with any number of cells as follows:

\[ X^2 = \sum \frac{(f_o - f_e)}{f_e} \]

Here \( f_o \) stands for observed frequency and \( f_e \) represents the expected frequency. The \( \Sigma \) sign means to add together the differences for all cells. For the present study, Chi-Square test was used to test the association between two categories wherever appropriate.

3.5.2. Analysis of Variance (ANOVA)

ANOVA is concerned with the analysis of variance. In a between-groups ANOVA, the focus of interest is on the variance between the groups. However, this between-groups variance can only be understood in the context of the variance that occurs within each group. In the analysis, a two-way Analysis of Variance (ANOVA) was done for comparing the challenges faced by different types of digital library initiatives.

This data analysis led to inferences and interpretations as presented in the subsequent chapters.