CHAPTER - VII

CONCLUSION
7.1 Introduction:-

There is tremendous marketing and peer pressure behind digital library products. These products are pushed by aggressive vendors. Similarly, prestigious research organizations have organized their own initiatives, and their reputations for innovation and leadership are inducing others to follow in this field. Many have a mixture of good and bad features. A close analysis of these points is carried out in the present research. The investigator has collected an extensive data and information on this topic. Since the nature of the research is conceptual, the efforts were made to scan the available resources on digitization, digital libraries and other related topics. After analyzing and interpreting the available information, the researcher has came out with following observations and findings. These observations are grouped in the following categories:-

A) Conservation and preservation of library materials
B) Digital library
C) Digitization
D) File formats
E) Copyright
F) Technological obsolescence
G) Scanning

7.2 Observations:-

A) Conservation and preservation of library materials

1. From the detailed study made regarding the preservation and conservation of library material, it is observed that most of the library material is deteriorated due to the excessive and inappropriate use and handling either by the library users or by the library staff.
2. The literary evidence shows that there is a lack of proper storage facilities, and adequate housekeeping practices regarding the preservation of library materials. The absence of facilities, materials and trained personnel in many libraries and the inadequate and uneven provision in others are major obstacles to the development of coherent programmes for the development of preservation and conservation of library materials.

3. Extremes or variations of temperature and relative humidity, light, dust and industrial pollution are important contributors to the destruction of library and archival materials, and their effects on documents and staff are still not properly estimated.

4. It is observed that the life expectancy of materials used in library and archival media varies greatly. The World Information report 1998 states that the lifespan of paper media is more as compared to electronic media. Many research works (Yale University Project) on durability of the storage medium has pointed out that microfilming is the best alternative method for the preservation and conservation of printed materials. Microfilming of historic document helps not only to protect the material in original form but also can enhances the accessibility for more than hundred years. Another notable point is that microfilm can be scanned more efficiently than the photographed originals.

5. It is observed that many library schools are providing education and training on library organization, classification, cataloguing and information retrieval but field of conservation and preservation is still neglected.

6. Digital preservation is a widespread problem in most of the institutions. It is observed that the institutions with digital preservation responsibilities are acquiring materials in digital form and creating digital files through conversion. It is revealed that institutions like New Zealand Digital Library,
Library of Congress and Pennsylvania University, Oxford University etc. have large collections of digital materials. These institutions cannot access some of their materials because they lack the operational or technical capacity to mount, read, or access files stored on some of the storage media in their holdings.

7. It is revealed that the field of film, photographs, audiovisual material and magnetic tape preservation is neglected as compared to paper media. The proper guidelines and well-established practices in conservation treatment are not available. Success in preservation of audiovisual material is highly dependent on information about its content. It is observed that media manufacturers are not willing to provide the information on their products.

8. It is found that there is a great problem in the preservation of magnetic tape as it is very difficult to find out which tapes will be the first to deteriorate beyond the point at which they can be salvaged.

9. Digital preservation policies and practices are not well developed in many institutions. These institutions do not have written policies for digital preservation. It is observed that common reason for not developing digital preservation policies is that they have not yet assumed responsibility for preserving materials in digital form.

10. Effective digital preservation requires life-cycle management of digital information from the point of creation through storage, migration, and providing access on a continuing basis. It reveals that the most common refreshing cycle is in the range of once in every three years to six years.

11. It is found that staff expertise, trained manpower is a major problem in institutions with digital preservation responsibilities. The lack of expertise
in digital preservation appears to be a significant obstacle to developing
digital preservation programs.

B) Digital library

12. By studying the existing literature (print and electronic) available on digital
library, it is clear that though the term 'digital library' is defined in more
than sixty different ways, still there is confusion about digital library.
Different scholars from different disciplines have given their own
explanations and justifications, which resulted in ambiguity and chaos.
The traditional library system has more systematic order as compared to
digital library.

13. To create a digital library one requires the same set of skills as building a
traditional library i.e. selection, preservation, arrangement, description,
and provision of access. Advanced technology offers the opportunity to
expand the scope of these skills to enter a new paradigm the global
electronic network. Most DLs are created to be a quick solutions to the
urgent community needs, hence planning of these systems for future
redeployment is not done properly.

14. Though many digital library projects were undertaken at national and
international level, only few of them are in operation. Those involved in
these projects themselves are confused about the digitization process. It is
found that many DLs are built in isolation as a response to the needs of a
particular community, in most cases, it is observed that personnel involved
in these projects have no prior experience.

15. Digital library software packages have been developed for organization
and retrieval of information. It is observed that out of these software
packages 'Contentdm' and 'Greenstone' are widely used. Greenstone is widely used on a large scale because it is available free of charge and can be downloaded from Internet.

16. The creation and maintenance of digital library is very expensive. The costs are incurred for production, for providing access and preservation. The cost to develop and operate a distributed architecture for long-term archiving, migration and backup of digital material at present is very high.

C) Digitization:-

17. Digitization is a critical issue, which involves a variety of processes; selection of material and methods to be adopted for conversion are very crucial among them. Regarding selection of materials, many organizations and universities have come up with several guidelines and standards. But many are not following these standards. Since the digitized materials of many organizations are not available on Internet, duplication of work is becoming a serious problem. Notable problems about digitization are variations in scanning resolution, different file formats and compression techniques, which resulted a hurdle in information acquisition and retrieval.

18. Different methods of converting the original material into digital formats are available at one’s disposal. Each one has its own advantages and limitations. Scanning can save the time and efforts of data entry, but cropping, editing the text and OCRing is quite cumbersome. To edit the text of particular subject, services of expert from relevant field is required. However it has been found that technically qualified people do most of the editing job. Digitizing a complex document (which has text, images, photographs, citations, handwritten information on single page) is very difficult. It is found that for effective retrieval of information, keying of the
data and information is the best solution as textual matter and images entered manually can be searched effectively. Shri Venkateshwara University, Tirupati is using scanning for preservation of the original manuscripts and then data is entered manually for providing better access.

19. It is cheaper to make microfilm and then scan it, than scanning from the original book. For printed books where black and white pages are the majority, and where a preservation copy is desired, it is recommended that microfilm be made and then scanned. Pages needing color or gray-scale can be separately scanned directly and then substituted for the bi-tonal images.

20. At present no specific standards are adopted for digitizing the library materials, like scanning resolution, digitizing software, file formats to store text and images etc. This makes accessing the material difficult.

21. It is observed that the current trend to reformat analog information into digital form is very high. The physical forms of film, photographs, and tape may soon be completely lost by the continuing process of converting the original material into digital form.

22. There are many difficulties of scanning the large size photographs, big size maps and oversized books, as the surface of scanners is predetermined.

23. Conversion of the textual information other than English (Roman) is very difficult as regional language OCR is not available. It is found that many government and non-government agencies are engaged in undertaking the research in this respect. C-DAC, Pune has developed a Devanagari OCR software called Chitrankan which is under test.
24. The quality of photographs, sound recording and visuals can be enhanced through digitization. There are many processes and editing techniques through which the dirty stains can be removed, fainted colours can be replaced with new colors etc.

25. The life span of the fragile material can be increased after digitization as original material is less handled because its surrogate is available to view and refer all the time.

26. Managing the large number of files is a tedious task because every page is saved as a separate file, naming them and remembering the content is very difficult. It is observed that scanning the page is very simple and easy, but linking these pages to form a full sized book is little difficult and requires special training in this respect.

27. It is observed that OCR is not suitable for old-fashioned text, calligraphic words and cursive writing. The common mistakes are found in recognizing the letters like B, O, R, C, e, c, p, f, etc. These produce as many as generally 30 to 40 errors on each page. It is observed that as an average about thirty minutes are required to edit a single page.

28. It is found that indexing and retrieval of photographs and illustrations is very difficult which makes it difficult to access these materials.

D) File Formats

29. Many file formats and storage media are used in the digital holdings of different institutions. Among these formats, the most common formats are image files, text files with mark-up, and ASCII files. It is also observed that most of these institutions have digital information in at least one of the
following formats: word processing files, audio video, and spreadsheets. It is also perceived that many institutions are maintaining digital information in several different formats. The number of file formats range from six to ten or more different formats.

30. The literary evidence shows that not a single file format satisfies the following requirements -

- efficient storage of image data in both compressed and uncompressed form,
- ability to access image data in a wide variety of formats,
- capability for standardized storage of image attribute information to facilitate image processing,
- extendibility for the accommodation of future needs, and
- reasonably wide levels of current acceptance.

31. By studying the various digitization projects, it is observed that TIFF (Tag Image File Format) is commonly used in most of these projects. This file format is very versatile, platform independent and open.

E) Copyright:-

32. Copyright is one of the most complicated issues that many librarians' are facing. Librarians who create digital archives are faced with even greater complexities. It is found that most of the librarians are not violating the copyright rules. It is observed that most of the material selected for digitization, is out of copyright.

F) Technological obsolescence:-

33. Basically, there are three ways in which digital materials can become inaccessible: (a) degradation of the media on which they are stored, (b)
obsolescence of software making it impossible to read digital files, and (c) introduction of new computer systems and peripherals that cannot handle older materials. Tapes and disks are all subject to physical decay and none of them has a lifespan that is comparable to that of preservation-standard microfilm or acid-free paper.

34. In spite of the relatively young age of digital collections in most institutions, they lack the operational and/or technical capacity to mount, read, and accesses some digital materials from their holdings. The most common storage media for which institutions lack access capability are floppy disks and open reel nine-track tape. There are also scattered problems with CD-ROMs, magneto-optical disks, DDS DAT tape, 3480 cartridges, and various audio and video formats.

35. Obsolescence of software and hardware leads to (partial) loss of information or functionality of files in their original format. Successive versions of programmes may be compatible, but software producers do not usually support compatibility over a long period. Programmes disappear from the market or can no longer be used on a new platform.

36. In the world of print, preservation can be achieved by preserving the paper object or creating a durable surrogate i.e. microform or other form. The equivalent in the digital world is to preserve a CD-ROM, or transfer its contents to another type of carrier. However, this does not achieve much more than preservation of the actual bits that make up a file. One cannot ensure that the information can be read and interpreted in the long run.

37. Through the digital library model presented herein, university libraries can bring all multimedia items on a particular topic together and provide better services to the users. The present model is based on MS-ACCESS software, which permits to bring text, pictures, audio/video clips and web
links together. Information on these items can be accessed on any of the fields created for this purpose. Librarians can insert scanned pages, downloaded text, sound recordings and video clips to build their own digital library as per the requirements of their users. University students, teachers and research scholars can benefit from this new tool to enhance, extend and support their teaching, learning and research activities.

G) SCANNING

38. Tonal (grayscale and color) images are necessary for preservation-quality reproduction of many manuscripts.

39. A spatial resolution of 300 dpi for tonal images is sufficient to capture all information in twentieth century materials like typescripts.

40. Some documents benefit from color reproduction but ambiguity remains about when color reproduction is necessary to serve researchers.

41. For preservation-quality images of routine documents—especially in the setting of a high-volume digitization effort—the benefit of reduced file size resulting from the modest application of lossy compression (at around 10:1 for JPEG) can be argued to outweigh the cost of the slight image degradation that results.

42. Access images can be produced by reprocessing the preservation-quality images. If end users are to be provided with easy navigation in the World Wide Web, "browser-capable" tonal display-access images must be provided. If end users are to be provided with clean printouts from a laser printer, separate binary print-access images should be considered; laser printouts from the reduced-size tonal display-access images will be less legible.
43. Objective tools for quality review are needed.

7.3 Recommendations:-

1. Items, which are very large by size for the flatbed scanner, require film intermediaries. Large items should be filmed on transparencies or microfiche. Digital camera should be used to capture these types of documents. Film intermediaries should be used for any artifacts or documents too fragile to put through the scanner, because heat of scanner can cause the erosion of ink.

2. Normal modern printed black and white text should be captured at 600 dpi in order to assure that all symbols, italic text, and other fine details are captured. Line art should be captured at 600 dpi if lines are fine and close together; if lines are bold and widely separated, lower resolution may suffice. Handwritten documents, typescripts, half tones, and similar materials should be captured at 300 dpi gray-scale. Printed documents such as maps and posters should be able to be captured fully at 200 dpi.

3. There is a great need to adopt standards and control tools for storage conditions, selection procedures for reformatting, and conservation treatment priorities to be adopted by libraries. It is urgently necessary to compile and distribute a set of guidelines, standards, and best practices for digital preservation.

4. Preservation of digital materials requires first of all definition and description of intellectual content. To be able to enforce the preservation principles in their daily routine, and to make sure the staff follows the guidelines and instructions, preservation knowledge of staff should be
brought up-to-date. To this end a training programme on preservation issues and training materials should be developed and produced.

5. As file formats and programs become outdated, preservation of digital materials has to deal not only with maintenance of the files themselves but also with ways of keeping them accessible. It is necessary to preserve the programs and make sure that they are well running on new platforms, or the files have to be converted to another format that can be interpreted by new programmes. It is necessary to use standards and open formats, adequate description and documentation, and the use of permanent names for online resources, which would facilitate long-term preservation and help to reduce costs. Producers of digital information should facilitate preservation efforts by using (official or de facto) standards. Emerging standards like XML and TIFF are promising because they are open standards not dependent on a specific platform.

6. A clear division of tasks and responsibilities, based on existing roles and expertise, needs to be established. It should be established how tasks can be shared between national heritage institutions and discipline-oriented organizations working for specific purpose. The leadership role in digital preservation of a number of heritage institutions worldwide should be acknowledged. Their pioneering work in exploring legal, organizational, technical and economic aspects can provide the basis for defining best practices which should be strongly promoted in the whole community.

7. Extensive training programmes are needed to preserve the digital information. Preservation of digital information requires new organizational structures, new approaches and new ways of thinking. Programmes will have to focus, not only on technical aspects, but also on training staff to deal with a changing environment and new directions. Personnel should be trained not only in the principles and practices of conservation and preservation, but also in the use and maintenance of available equipment.
8. The success of digitization and digital library projects depends on the transformation of these projects into services, working standards and extensible, adaptable tools for future use. In order to produce high profile, well-used services there must be dedicated funding, sponsorship, and education grants. Government organizations and other funding agencies like UGC, National Library, CSIR, ICSSR, INFLIBNET etc. should take a lead for funding the various projects.

9. It is strongly recommended that those who are digitizing their materials should make it available through their website so that duplication of work can be avoided. A national portal for digital material should be created.

7.4 Suggestion for future studies

Research is an ongoing activity. The positive results of today may prove negative tomorrow. The limitations and disadvantages of any present system can be eliminated by systematic study and research. Digitization and digital libraries have many dimension and new vistas, further research needed to uncover its manifold benefits to entire humanity. The researcher would like to suggest following topics for further study:

a) Digital world is moving ahead with great speed. New means and methods are developed to keep track of the digital information. Many software vendors are developing digital library software packages to handle and retrieve digital information. Further research can be undertaken to explore these software packages with their merits, demerits and limitations if any. This will help to select suitable digital library software for acquisition and retrieval of digital information.
b) Large number of scientists and researchers are using digital information to carry out their projects and research works. It is necessary to find out the information seeking behavior of these scientists in digital world. Is this new media is really helping them to satisfying their needs.  

7.5 Conclusion

As Shakespeare so eloquently notes in the couplet of the famous 18th sonnet, the printed word has a kind of immortality that few other things can claim. The word “this” in the final line refers to the sonnet itself, thereby proving its own point.

Shall I compare thee to a Summers day?
Thou art more lovely and more temperate:
Rough windes do shake the darling buds of Maie,
And Sommers lease hath all too short a date:
Sometime too hot the eye of heaven shines,
And often is his gold complexion dimm'd,
And every faire from faire some-time declines,
By chance, or natures changing course untrim'd:
But thy eternall Sommer shall not fade,
Nor loose possession of that faire thou ow'st,
Nor shall death brag thou wand'r'st in his shade,
When in eternall lines to time thou grow'st,
So long as men can breath or eyes can see,
So long lives this, and this gives life to thee.

A modern, digital version of the couplet would have to be something like the following:

So long as the magnetic flux on this disk has not been disturbed, 
and so long as humans retain the appropriate size and speed disk drives, 
and so long as they have hardware controllers and software device drivers capable of reading the bits from this disk, 
and so long as they have access to the software that encoded the file structure, 
and character codes employed in the bit stream of this document, 
and so long as they can still find or recreate the computing environment necessary to run that software, 
and so long as they can still breathe or see, 
So long lives this,...