6. FINDINGS AND SUGGESTIONS

6.1 Introduction:

The advancements in ICT and their proper utilization by research and academic librarians are not only strengthening the capabilities of libraries but also helping to achieve excellence. The present study is about the backup practices in digital library of research and higher learning institutes in Gujarat. This study aims to get information about backup devices, services available for backup, post backup practices, preservation of backup, to identify obstacles occurring to it and human resources & their skill upgradation for baking up etc. For this purpose, questionnaire was designed (Appendix 1) and distributed among 12 selected libraries of research and higher learning institutes and of which 10 were received back.

Data collected from libraries have been analysed and presented in Chapter 5: Data Analysis and Interpretation. However, based on this findings, suggestions and areas of further research have been derived and same have been presented below under respective headings.
6.2 Findings:

From the detailed analysis and interpretation about the backup practices in digital library of research and higher learning institutes in Gujarat, the findings are summarised below.

**Digital Resources/Institutional Repository /Archive:**

- Eight (80%) libraries have created institutional repository and one (10%) library has developed digital archive for management of their digital content.
- Only one (10%) library has not developed digital repository /archive for its digital content. This library has only digital resources.

**Software used for creation of Digital Repository /Archive:**

- Libraries are using Open Source Software for creation of Digital Repository /Archive. Seven (70%) libraries are using Dspace. Two (20%) libraries are using GSDL. None of library uses E-print.
- Dspace is more popular among all digital library/ institutional repository software.

**Necessity of Backup:**

- All libraries (100%) know the importance of backup and considered backup is necessary.

**Reasons for Backup:**

- Libraries gave equal response to digital preservation (40%) and safety purpose (40%).
- 20% libraries take backup for disaster recovery.
Data loss experience:

- Two (20%) libraries have faced data loss experiences. They overcome this with the help of IT professional.
- Eight (80%) libraries do not have data loss experience.

Backup Policy:

- Six (60%) libraries do not have backup policy
- Four (40%) libraries have backup policy. But these libraries neither provide any detail pertaining to written policy nor actual policy on paper.
- There is lack of awareness pertaining to backup policy among libraries.

Budget for Backup:

- Nine (90%) libraries do not have special budget for backup.
- Only one (10%) library has provision for special budget for backup.

Criteria for data selection for backup:

- Majority of libraries consider criteria such as content value, archival value, frequency of change, rare data etc. for data selection for baking up.
- Cultural heritage value receives least response among all criteria.
- Currency/recency of content also taken care at time of data selection.

Frequency of Backup:

- Libraries take backup of digital content at regular time interval such as on the spot, daily, weekly, fortnightly, monthly & annually etc. based on their requirement, convenience and policy.
Four libraries do backup of their content on the spot, three libraries carry out backup weekly and monthly, two libraries fortnightly and only one library performs backup daily and annually.

It is further found that only one library carries out backup daily, monthly and annually.

Two libraries do backup two times i.e. on the spot & fortnightly and on the spot & monthly.

Seven libraries carry out backup of their data only once.

Backup Method:

All libraries (100%) performed full backup of their content manually.

Only two (20%) libraries carry out incremental backup besides full backup.

Only one (10%) library performs backup manually as well as automatic.

Very few libraries perform backup using multiple method.

Backup devices used:

Most of the libraries take backup using multiple backup devices.

Hard Disk is popular backup device among all backup devices. All libraries use it.

Three libraries perform backup on NAS and another machine or system.

DVD, CD and magnetic tape are least use.

None of the libraries prefer flash memory & Blue ray disk as backup media.
Criteria for selecting Backup devices:

- Majority of libraries give response to security, capacity and speed of backup device respectively.
- Price is least concerned for this purpose.
- Besides these criteria, other two criteria have taken into consideration by libraries i.e. availability and currency of backup device.

Online backup services:

- Two (20%) libraries use online backup service.
- Among this libraries, one library opt free backup service i.e. Dropbox and another library use proprietary online backup service i.e. www.box.net. This site takes backup annually.

Backup storage locations:

- Libraries stored backup devices in fire proof cabinet (30%), at secure offsite location (30%) and physical data storage vault (20%) etc.
- Two (20%) libraries keep backup devices near the computer.
- None of library staff stores backup devices at home or any other place.

Post backup Practices:

- In post backup practices, labelling get highest response (07), followed by preservation (05), classification (02) verification (02), restoration (02) and authentication (01).
- None of the libraries have performed all post backup practices.
Preservation of backup:

- All libraries (10) use replication as preservation strategy.
- Seven libraries store backup devices in secure & proper environment.
- Four libraries use migration/refreshment.
- Two libraries do cleaning of storage media and perform auditing for preservation purpose.
- None of the libraries use all strategies.
- Libraries use multiple preservation strategies.

Obstacles in present backup practices:

- Five (50%) libraries face different obstacles in backup practices such as time consuming process, backup device requires special care in handling, backup-not simple & easy as sometime system did not support/compatible, frequent technological change, fast hardware obsolescence, lack of funds and lack of technical skills etc.

Data selection for backup done by:

- Data selection for backup done by system administrator (40%) followed by librarian (30%) and other i.e. library staff (30%).

Backup taken by:

- In seven (70%) libraries, library staffs take backup.
- In remaining libraries, system administrator/IT manager (20%) and librarian (10%) take backup.
Detail of person concerned for backup:

- Eight persons have knowledge of library and information science and two persons have knowledge of IT.
- Only two libraries have person with knowledge of both field i.e. LISc and IT.
- All persons have different designation.

Hire IT Professional for Backup:

- None of the libraries hire IT professional for backup.
- All libraries perform backup with the help of staff available in library and IT department of respective institutes.

Training for Backup:

(a) Training for taking backup:

- Seven (70%) libraries are providing training to concerned staff for backing up.

(b) Training for Latest backup Technology:

- Eight (80%) libraries did not provide training pertaining to latest backup technology.
- Two (20%) libraries keep concerned staff up to date with latest backup technology.
6.3 Hypotheses testing:

Following are the hypotheses proposed and their findings:

**Hypothesis: 1 Majority of libraries have faced data loss experience.**

Table: 5.5 is related to data loss experience. It reveals that 20% libraries have faced data loss experiences. Therefore the hypothesis stands partially confirmed.

**Hypothesis: 2 All libraries have backup policy.**

Backup policy is essential for any digital project to be successful. Table: 5.6 reveals that 60% libraries do not have backup policy and 40% libraries have backup policy but not in written form. Thus this hypothesis is rejected.

**Hypothesis: 3 Multiple backup methods are practiced in almost all libraries.**

In Table: 5.10 reveal that only two libraries carry out full and incremental backup. Only one library performs backup manually as well as automatic. All the libraries practice full manual backup. So this hypothesis is partially rejected.

**Hypothesis: 4 In majority of libraries, multiple backup devices are used.**

Table: 5.11 reveals that most of the libraries use multiple backup devices. Thus the hypothesis is accepted.

**Hypothesis: 5 In most of the libraries, online backup services are used.**
Table: 5.13 reveals that 20% libraries take advantage of online backup services. Thus the hypothesis stands partially confirmed.

**Hypothesis: 6 Almost all libraries are face problems in taking backup.**

Table: 5.17 reveals that 50% libraries face different problems in backing up data. Thus the hypothesis stands partially confirmed.

**Hypothesis: 7 All libraries are provide training to concerned person about latest backup technology.**

Table: 5.22(b) reveals that 20% libraries are providing training to concerned person about latest backup technology. Therefore, this hypothesis is partially rejected.
6.4 Fulfilment of Objectives:

1. **To know backup practices in libraries.**

   Questionnaire was designed to get information regarding backup practices in libraries of research and higher learning institutes. Analysis of collected data in *Chapter: 5* provides information about backup practices in libraries under the study. It was seen that all libraries have backup practices in less or more extent.

2. **To identify backup devices and services (or backup tools) used by libraries.**

   Table: 5.11 & 5.13 are about backup devices and services. Libraries use various backup devices such as Hard Disk, DVD, CD, magnetic tape & NAS etc. Hard Disk is popular among all backup devices. Most of the libraries take backup using multiple backup devices.

   Out of 10, only two libraries prefer online backup service such as free backup service i.e. Dropbox and proprietary online backup service i.e. www.box.net.

3. **To find out obstacles in taking backup.**

   Table: 5.17 provide information about the problem faced by the libraries during backup. It is observed that 50% libraries face obstacles in baking up such as time consuming process, backup device requires special care in handling, backup-not simple & easy as sometime system did not support/compatible, frequent technological change, fast hardware obsolescence, lack of funds and lack of technical skills etc.
4. To suggest standard backup practices to libraries.

Based on research, researcher has prepared standard backup practices.

**Standard Backup Practices**

- Prepare backup plan at initial stage of Digitization Project / Digital Library/ Digital Archive or Institutional Repository.

- Create committee for backup plan.

- Create a backup policy that clearly identifies:
  
  - Roles & Responsibilities (Select dedicated skilled staff for backup. Provide training regarding backup & latest backup technology)

- Data selection for backup

- Backup method such as Full backup, Incremental backup, Differential backup manually as well as automatic.

- Backup devices such as data tapes, hard disk drives, solid-state drives, optical disks etc.

- Backup software

- Online backup services: Use various online backup services such as Dura Cloud, Google Docs., Dropbox, Sky Drive, Amazon cloud service - Amazon S3 (Simple Storage Service), Box.net, ADrive etc.

- Frequency of backup like on the spot, daily, weekly, fortnightly, monthly & annually etc.

- Recommended file formats to be used such as pdf, tiff, JPEG 2000 etc. Use non-proprietary, standard formats
✓ Metadata

✓ Post backup practices: Tagging / Labelling, Classification, Verification, Authentication, Restoration, Preservation etc.

✓ Backup storage locations: In fire proof cabinet in the room, in a physical data storage vault, secure offsite location, any other etc.

✓ Preservation of backup: Use multiple preservation strategies such as store in secure & proper environment, auditing, migration/refreshment, replication, cleaning of storage media etc.

➢ Versioning

Use consecutive numbers and letters to keep track with changes to a file throughout continuous editing & revisions. This will help in quickly differentiate between files with similar names.

➢ File Naming

Use file names that are consistent, descriptive, and concise so that one can quickly find out and identify the file at a later time.

➢ Review backup policy and plan periodically.

➢ Minimize or remove reliance on others/user to perform own manual backups (if possible)

  ▪ Implement standardized and automatic backups

  ▪ If possible, put experts in charge of this task (computer staff) as they are more likely to keep up-to-date regarding software updates, hardware issues, best practices, etc.
Check backups manually

- Most the backup software will have a log file that contains details of the backup (which files, when the backup was created)
- However, don’t rely solely on the log file
  - Even if a log file states the backup was successful, you still need to check the backup to make sure the files are there and accessible
  - Look at file dates and file sizes to see if they match; calculate a checksum on the original and archived file and make sure they match
  - Ensure *Usability, Authenticity, Discoverability & Accessibility* of digital content that backups and archives earlier.

Keep backups of the backups

- Necessary for high-value data
- Usually take multiple backups using multiple format, multiple backup devices & multiple backup method and store multiple locations

Proper backup practices ensure continues functioning, increase life expectancy of data & minimizes the risk of data loss.
5. To design backup policy for libraries.

On the basis of the present research and literature review, researcher has designed backup policy for libraries.

**Backup Policy**

**Purpose:**

The purpose of this policy is to establish standards and guidelines for the backup process within the library. Proper backup is necessary for the library to achieve its objectives efficiently. Library staff must protect *Usability, Authenticity, and Discoverability & Accessibility* of digital content.

Effective implementation of this policy will

- Safeguard the information assets of library
- Prevent loss of data in case of any disaster/accidental deletion of data/corruption of data or system failure.

System administrator or librarian will be responsible for developing a policy for backing up the systems depending upon configuration, software applications and nature of data and other factors. It should be reviewed and revised regularly. This policy must be documented and made available to stakeholders for implementation.

Administrator will ensure the procedures followed strictly and implemented as per rules.

Library should maintain backup infrastructure, including upgrading the hardware & software as needed.
Scope:

The policy should apply to all the servers and computers operating at library.

It also applies to all staff, student & 3rd parties who use IT services connected to the library network and who process or store information acquired by library. All stakeholders are responsible for arranging adequate backup procedures for data held on IT devices assigned to them.

Strategy:

- Standard backup practices must be performed and monitored regularly.
- Library must take multiple backups using multiple formats, multiple devices, and multiple methods and store it at multiple locations.
- All post backup practices such as Tagging / Labeling, Classification, Verification, Authentication, Restoration, Preservation etc. should be performed properly, without fail.
- Multiple preservation strategies such as store in secure & proper environment, auditing, migration/refreshment, replication, cleaning of storage media etc. should be used based on requirement to ensure access over a period.
Procedures:

All backup must conform to the following best practice procedures:

1. Select data considering criteria such as content value, archival value, frequency of change, rare data etc. for backing up.

2. Back up all data, operating systems and utility files adequately and systematically.

3. Back up records of software licensing.

4. Ensure backup media /devices such as CD, DVD, and Hard Disk & Tape Drives etc. for necessary backup.

5. Set infrastructure for taking backup over the network for sub-sections.

6. Use remote backup services through secure & trusted server on the intranet.

7. Backup schedule:

   - Incremental backup - Daily - Monday to Friday and data locate on-site
   - Full backup - Weekly - Saturday and data locate offsite
   - Daily backup take place on five days rotation
   - Weekly backup take place on a week rotation
   - A monthly backup occurs the last calendar day of the month and are on a twelve month rotation.
   - All stakeholders backup their data before updating or upgrading software on their computer.
Special backups need for longer retention periods during special situation such as system upgrades and major projects.

8. Carry out multiple backups using multiple formats, multiple devices, and multiple methods and store multiple locations.

9. Perform all post backup practice:

- Label backup media precisely and maintain accurate records identifying criteria such as: system name, creation date, backup period, backup set name, concern person detail/contact number etc.

- Store backup copies in environmentally protected and access controlled secure offsite location. Store daily backup onsite in physically secured disaster proof safe location in building separate from server room. Provide access to authorized people.

- Preserved copies make available on authorized request.

- Maintain record of the physical movement of all backup copies.

- Maintain tape drives, hard disk & other backup media according to manufacturer’s recommendations.

- Do not leave media unattended, during transport or changes of media.

- Verify backup periodically to ensure the integrity of the operation.
➤ Check retention of backup at regular interval.

10. Use multiple preservation strategies such as auditing, migration/refreshment, replication, cleaning of storage media etc. regularly based on requirement to ensure access over a period.

11. Regularly maintain log of backup & recovery process for each system and review it.

Enforcement:

If any employee found to violate this policy should be subject to disciplinary action as per the rules of the parent body/organisation.

Responsibility:

It is the responsibility of the librarian/System administrator to follow proper backup practices in library. There must be another nominated person who can take care of the backup in absence of the librarian/system administrator.
6.5 Suggestions:

Major suggestions, based on study are given below:

- Library should have well defined backup plan for digital content/digital library/archive or institutional repository. And it should be reviewed regularly.
- Backup policy should be formulated at time of undertaking the digital library projects.
- There should be special provision for backup in budget.
- Libraries should take multiple backup in multiple formats on multiple backup devices and it should be stored at multiple locations.
- Backup should be performed using different backup methods manually as well as automatic.
- Post backup practices should be performed.
- Libraries should use multiple preservation strategies.
- Retention of backup should be checked regularly.
- Libraries should take benefit of various online backup services.
- There should be dedicated skilled staff for backup. They should get training regarding latest backup technology.
- There should be Co-Operative Backup Centre like Data Centre at state or national level where all libraries can store their digital content and restore from it, in case of any disaster. This centre or libraries can tie up with BSNL or IT Company.
The country should have a Backup Policy for library at national level. A committee should be formed by Ministry of Communication and Information Technology to prepare a white paper on the subject and draw implementation strategies for various stakeholders.
6.6 Suggestions for further research:

The possible further areas of research can be:

- Use of online backup services in libraries.
- Feasibility study of creation of data backup centre for libraries.
- Life span of various backup devices use in libraries.
- Evaluation of various online backup services and software.
- Challenges in digital libraries with special reference to preservation.
- Skill and competencies required for the digital preservation pertaining to backup practices.
- Safeguards for digital preservation.