2. Review of Literature

2.1 Literature Review:

2.1.1 Concept and Meaning:

A literature review discusses published & unpublished information in a particular subject area, and sometimes information in a particular subject area within a certain time period.

A literature review can be just a simple summary of the sources, but it usually has an organizational pattern and combines both summary and synthesis. A summary is a recap of the important information of the source, but a synthesis is a reorganization or a reshuffling of that information.

It might give a new interpretation of old material or combine new with old interpretations. Or it might trace the intellectual progression of the field, including major debates. And depending on the situation, the literature review may evaluate the sources and advise the reader on the most pertinent or relevant.

The format of a review of literature may vary from discipline to discipline and from assignment to assignment.

A review may be a self-contained unit an end in itself or a preface to and rationale for engaging in primary research. A review is a required part of grant and research proposals and often a chapter in theses and dissertations.
Generally, the purpose of a review is to analyze critically a segment of a published body of knowledge through summary, classification, and comparison of prior research studies, reviews of literature, and theoretical articles.

A literature review is the effective evaluation of selected documents on a research topic.

The evaluation of the literature leads logically to the research question.

A ‘good’ literature review…..

- is a synthesis of available research
- is a critical evaluation
- has appropriate breadth and depth
- has clarity and conciseness
- uses rigorous and consistent methods

Literature review provides handy guide to a particular topic. If researcher has limited time to conduct research, literature review can give an overview or act as a stepping stone.

The purpose of a literature review is to take a critical look at the literature (facts and views) that already exists in the area of research topic.

A literature review is not a shopping list of everything that exists, but a critical analysis that shows an evaluation of the existing literature and a relationship between the different works. It demonstrates the relevance of the research.
Literature can include books, journal articles, internet (electronic journals), newspapers, magazines, theses and dissertations, conference proceedings, reports, and documentaries.

2.1.2 Importance of literature review in research:

Gall, Borg, and Gall (1996) argue that the literature review plays a role in:

- Delimiting the research problem,
- Seeking new lines of inquiry,
- Avoiding fruitless approaches,
- Gaining methodological insights,
- Identifying recommendations for further research,
- Seeking support for grounded theory.

Hart (1998) contributes additional reasons for reviewing the literature, including:

- Distinguishing what has been done from what needs to be done,
- Discovering important variables relevant to the topic,
- Synthesizing and gaining a new perspective,
- Identifying relationships between ideas and practices,
- Establishing the context of the topic or problem,
- Rationalizing the significance of the problem,
- Enhancing and acquiring the subject vocabulary,
- Understanding the structure of the subject,
- Relating ideas and theory to applications,
Identifying the main methodologies and research

Techniques that have been used and

Placing the research in a historical context to show familiarity with state-of-the-art developments.

Another purpose for writing a literature review is that it provides a framework for relating new findings to previous findings in the discussion section of a dissertation. Without establishing the state of the previous research, it is impossible to establish how the new research advances the previous research.

2.1.3 The whole process of reviewing includes:

a. Searching for literature

b. Sorting and prioritizing the retrieved literature

c. Analytical reading of papers

d. Evaluative reading of papers

e. Comparison across studies

f. Organizing the content

g. Writing the review

Here researcher referred various online and print resources such as websites, articles, books, presentation, conference proceedings and thesis etc. regarding digital preservation (need for digital preservation, definition & meaning of digital preservation, preservation strategies, preservation initiatives), backup (backup devices & its issues, backup
methods, backup process, backup policy) and studies related to digital preservation & backup at national and international level.

**Review of Relevant Literature:**

Application of Digital Technology in creation, storage & dissemination of information and various digital libraries initiatives at national & international level, results in digital libraries. Libraries at global level face issues of digital content. Libraries are more dependent on technology. We know that books printed before 25 or 50 years back can be read till today. We cannot guarantee the same for digital content created recently. The main reason is the dependency on equipment & technology (Arms, 2000) and fast changes in storage media.

### 2.2 Digital Preservation

#### 2.2.1 Need for digital preservation:

There is an inverse relationship between capacity & longevity of the media used to store the information (Paul Conway, 1996). The rate of change in computing technologies is such that information can be rendered inaccessible within a decade (Julian Jackson). The average life of a URL and e-mail is 44 days (Kahle, 1996). Because there is such a large wealth of information that is now digitized, it is almost impossible for libraries and archives to save it all. Up to 80% of websites are updated and/or gone within one year of creation; this makes it nearly impossible for professionals to keep up (Lasfargues, Martin & Medjkoune, 2012). Even we also
experienced that information available online, sometimes disappeared within a fraction of second.

Jain S. and Dr. Sanjeev (2012) in their respective article highlights the digital information is very dynamic. The databases are always being updated. What one gets on the website today now may not be there in the next few seconds. The digital media are fragile with a limited shelf life. Further still, the digital information on the storage devices with time will be rendered unreadable by obsolescence of technology, this due to the fact that information technology evolves very fast and the old systems are no longer in the use. To preserve the digital information, libraries will keep on migrating information from one digital hardware & software configuration to subsequent. The policy should address these issues. If preservation of digital resources will not be in place then future generation will look back at this time as a digital dark age a time when, somehow, the records of human knowledge went missing.

Li, Y. (2011) points out digital information are fragile and face many threats including technological obsolescence and deterioration of digital storage media. Smit, Hoeven and Giaretta (2011) discuss some people seem to think that in a digital world, digital content is easy to store at somewhere digitally due to misperception that “as long as it is digital, it is safe”. But it is not so easy. Digital content is not in usable condition due to technological; format, hardware, software obsolescence or due to bit-rot. In short anything digital is fragile and susceptible to decay. As amount
of digital content increasing digital preservation is becoming an ever-pressing issue. The prospect of losing digital content is therefore highly alarming; a Digital Dark Age may be looming. The term digital Dark Age was coined by Terry Kuny at IFLA conference (1997) and applies to a possible future situation where it will be difficult or impossible to read historical documents because of inadequate digital preservation.

Blackeslee, S. (1990) in article “Lost on earth - wealth of data found in space” states the record of 1976 Viking landings on Mars were unprocessed. When later analyzed, the data were unreadable as they were in an unknown format and the original programmers had either died or left the organization. Computer systems and software application change so rapidly that there is no guarantee that the existing data resources will be accessible and usable on future computing platforms or software versions (Wilson, 2007).

Software & hardware obsolesce are great technical threat for digital content.

Lawrence et al. (2000), Granger (2000), Rosenthal et al. (2005) and Glandney (2007) all observed that digital objects will cease to be accessible without active management and intervention.

Rath and Samantaray (2008) indicate that preservation & management of digital resources are concerned with 3 types of problems such as;

1. Preservation of artifact or medium on which the information is stored (e.g. tape disk, optical disk, CD) and preservation media,
2. Preservation of software and storage formats, which can become obsolete and create access problem (technology preservation) and

3. Preservation of the information content from being corrupted from its original form (intellectual preservation).

Further they argue that, many of the libraries and information centers have given a serious thought and effort to building digital collections but very little thought to preserve these digital resources.

In this direction, British Library (2013) opines that preservation of digital content is not straight forward. It requires action and intervention throughout the lifecycle, far earlier and more frequently than does physical collection.

Hence, preservation of collection is fundamental for success of any library: if library cannot preserve them, it will fail in core mission to enable access for future generations. It is very true for digital collections, too. Yet over the very long term, the inherent instability and transient nature of digital content makes preservation a significant challenge, one faced not just by the library but by institutions worldwide whether it may be private, research, scientific, educational, commercial and other government bodies.

Hedstrom (1998) in their article says that digital preservation is constrained by protocols, and proven methods for preserving digital information and by the tendency to consider preservation issues only at the end of a project or after a sensational loss. With few exceptions, digital library research has focused on
architecture and system for information organization, retrieval, presentation and visualization and on the administration of intellectual property rights (Levy and Marshall, 1995). The critical role of digital libraries and archives in ensuring the future accessibility of information with enduring value has taken a back seat to enhancing access to current actively used material. As a consequence digital preservation remains largely experimental and replete with the risks associated with untested methods.

Preserving digital content is challenging work. According to Ricksarkivet (National Archives) in Sweden, the costs of digital storage are much higher than generally believed (Palm, 2006).

There are many challenges in digital preservation and solution for its are in their infancy (Watson, 2009).

2.2.2 Definition & Meaning of Digital Preservation:

According to the association for information & image management, digital preservation is the ability to keep digital documents and files available for time periods that can transcend technological advances without concern for alteration or loss of readability.

Whereas RLG/OCLC report defines digital preservation as series of managed activities necessary to ensure continued access to and preservation of digital material.
Glossary of LIFE explains digital preservation is a process of ensuring that a digital object is accessible over the long term.

According to Hixson (2004), preserving digital content entails far more than making backup copies and storing them in disparate locations. Digital preservation is to extend the usable life of machine readable files and protect them from media failure, physical loss, and hardware & software obsolescence. The activities include:

(1) Ensuring the long term maintenance of bit stream (the zeros and ones):
   Baking up files and keeping a copy at an offsite location.
   Running checks to track the deterioration of storage media, files & bit streams.

(2) Providing continued accessibility of the contents:
   Viability-making sure that information is readable from the storage media:
   Readability- making sure that information is viewable by humans and able to be processed by computers.

(3) Understandability- making sure that information is able to be interpreted by humans.

Smit, Hoeven and Giaretta (2011) indicate that there is a quite a bit of confusion as to what the concept of digital preservation actually entails. Digitization, archiving, long term storage, or open access are not digital preservation. According to the definition as used in Open Archival Information System (OAIS) and by Giaretta, two key elements are central in digital preservation-usability and understandability:
authenticity over time of a digital object (Being able to reuse, understand and authenticity). Digital preservation is above all and more.

In other words preservation makes no sense if the preserved objects cannot be made retrievable, discoverable, accessible & reusable, and if they cannot at the same time provide assurance that they are what they claim to be.

Digital preservation is the act of physically & intellectually protecting and technically stabilizing the transmission of the content and context of electronic records across space and time, in order to produce copies of those records that people can reasonably judge to be authentic (Hitchcock et al. 2007).

Research Libraries Group and Online Computer Library Centre (2002) defined digital preservation as the activities necessary for ensuring both the long term maintenance of a byte stream and continued accessibility to its contents.

2.2.3 Preservation Strategies:


1. **Investment digital preservation strategies** can be adopted at the initial stage of a digital object. It includes **use of standards**- use of preferable open, widely available, supported or agreed standards and file formats that ensure stability and long term support. PDF/A is widely used standard for long term preservation of document (Strodl et al. 2007). RTF (Rich text format) is also used.
**Normalization**: converts digital content into one or more preferred formats. Example: JPEG/GIF image converts into TIFF and word document into Open Document Text (ODT).

**Encapsulation**: means of binding together data & the means of providing access to it. Boudrez (2005) argues that encapsulation is a storage technique in which metadata is kept together with digital content. Which are expected to remain unexploited for long periods of time. All changes to files, once created, must be documented and embedded into the metadata in order to preserve the integrity of the information (Gracy & Kahn, 2012). The US National Archives and Records Administration (NARA) holds that if something is worth preserving digitally, it must be preserved as close as possible to its original state (Galloway, 2009).

Ferreira opines this strategy is for digital content.

**(2) Short term Digital Preservation Strategy:** This is for 1 to 5 years and covers technology preservation, refreshing.

**Technology Preservation**- Preservation of the original technology use to create the digital content in order to preserve the functionality and “look & feel” of the content. Strategy stresses on the technological environment than content. Thibodeau (2002) argues that this might offer a short term solution. It is not viable for long term.
Refreshing: Mitigates media degradation and obsolescence. Refreshing is copying digital content from storage medium to another. Reis and Lindley (2007) say that media migration is the transfer of data from one storage medium to another storage medium.

Andrews and Law (2004) report that refreshing cannot be consider as a solution of digital preservation as it does not cover the transfer of the entire digital environment, but only the physical storage medium.

(3) Long term Digital Preservation Strategy:

Migration and emulation are widely used long term preservation strategy for digital content.

Migration- Transferring data to latest platform. According to TADI, migration is the periodic transfer of digital content from one hardware/software configuration to another or from one generation of computer technology to a subsequent generation.

Migration stresses on the digital content, rather than its environment. According to Reis & Lindley (2007), aim of migration is to change digital content in such way that hardware & software developments will not affect its accessibility. This covers content migration from a source format in to a target format, and media migration from one digital medium to another medium.
Migration has risks. Granger (2000) observed the notable danger of migration is that of data loss or in some cases, the loss of original functionality or the look & feel of original platform.

**Emulation:**

In emulation original hardware & software & operating system are stored along with digital content. Current technology is used to mimic the original environment. Granger (2000) reports that the purpose behind emulation is to be able to access or run original data /software on current platform that emulates the original platform. It requires detail information of the original environment on which emulation is to base in future.

Lalthlanthangi (2012) says that there are several solutions for long term access such as conversion, refreshing, migration, emulation etc.

OCLC (2006) developed 4 points strategy for the long term preservation of digital content.

There are several other strategies used for preservation of digital content. Such as migration, formatting, emulation, universal virtual computer, encapsulation, digital archive, data archaeology, preservation of used technology and metadata etc. According to Smit, Hoeven & Giaretta(2011), most of the strategies take a strong IT perspective such as normalization/standardization, refreshing, transformation (migration), emulation, semantic metadating, combination of these strategies.
Considering environment of organization and infrastructure, digital preservation strategies are more concerned with policy, economic and social aspects such as common standards for metadata, identifiers, certification and auditing, economic sustainability, holistic approach, future citability and rights managements, proper training of preservation etc.

They suggest sensible digital preservation strategy which covers combination of different technical strategies, network of trustworthy deposit places and archive, involvement of all stakeholder, infrastructure with inter operable identifiers.

Luan F. & Nygard M. (2010) give taxonomy of preservation strategies. In which auditing, refreshment, replication and federation etc. strategies are for storage accessibility maintenance. While for bits manipulability maintenance; computer museum, emulation, encapsulation, universal virtual computer, batch migration, migration on access etc. strategies are covered.

Maharana (2001) argues libraries are facing issues of digital content at global level. Although, libraries have experimenting with various preservation strategies, still there is a need of a technological feasible, financially affordable and widely acceptable strategy for digital preservation.

Kahn M. (2004) suggests strategies such as copying, reformatting or conversion, migration, emulation etc. for long term preservation.

Li, Y. (2011) reported that libraries are backing up digital content and stored it in a secure storage system. Most of libraries had checksum algorithm to detect errors
in the data stored in IR. Few libraries use digital preservation strategies such as migration, emulation and refreshing. LOCKSS, MetaArchive, Dura Cloud, iRODs, CDL curation services etc. system used to support digital preservation.

According to Rath P. and Samantaray M. (2008) technology preservation, media refreshment, migration, normalization, emulation etc. strategies are being explored to deal with technological obsolescence.

2.2.4 Preservation Initiatives:

Kirchhoff (2008) says the research of the future requires access to the research of the past. This access cannot be assured without reliable long-term preservation of scholarly digital content. Near-term access can be guaranteed with backup and access system redundancy. Mid-term access can be protected with byte replication. But assurance of long-term access requires digital preservation – the series of management policies and activities necessary to ensure the enduring usability, authenticity, discoverability, and accessibility of content over the very long term. And also discuss in brief guidelines available for digital preservation such as

- PREMIS (Preservation Metadata: Implementation Strategies)
- TRAC (Trustworthy Repositories Audit Certification: Criteria and Checklist)
Gaur & Tripathi (2012) highlights various problems of digital content and elaborates how digital preservation of e-resources is more demanding and challenging than preserving print copies of journals. It also gives a bird’s eye view of various projects initiated for archiving digital content of scholarly journals such as

- Portico
- LOCKSS (Lots of Copies Keep Stuff Safe)
- CLOCKSS (Controlled Lots of Copies Keep Stuff Safe)
- PANDORA (Preserving and Accessing Networked Documentary Resources of Australia)
- KOPAL
- Pub Med Central
- Ohio Link
- e-Depot System
- National Digital Information Infrastructure and Preservation Programme (NDIIPP-LC)
- Data-PASS
- Data Preservation Alliance for the Social Sciences
British library (2013) discusses their four preservation strategies (1) Ensure digital repository can store and preserve their collections for the long term (2) Manage the risks and challenges associated with digital preservation throughout the digital collection content lifecycle (3) Embed digital sustainability as an organizational principle for digital library planning and development (4) Benefit from collaboration with other national and international institutions on digital preservation initiatives.

And mentions some preservation projects;

- EU-project PARSE. Insight
- Portico, CLOCKSS (US)
- PLANETS project (2006-2010)
- SCAPE project (2011-2014)
- LIFE (2005-2010)
- AQUA (2011)
- APARSEN (2011-2015)
- SPRUCE (2012-2013)

Luan and Nygard (2010) discuss preservation of digital information is difficult because of both technology as well as relevant organizational context change over time. Without careful protection, there is real risk in retrieval and usability of digital content. This article gives an overview of the current main preservation system;
- LOCKSS
- Eprints
- Dspace
- e-Depot
- FEDORA (Flexible Extensible Digital Repository Architecture)
- iRODS (Integrated Rule Oriented Data System)

Also provide information regarding current research groups engaged in long term preservation of digital content;

- PADI (Preservation Access to Digital Information) (Australia)
- Inter PARES (International Research on Permanent Authentic Records in Electronic Systems)
- DCC (Digital Curator Centre)
- CAMiLEON (Creative Archiving at Michigan & Leeds: Emulating the Old on the New)
- PLANETS (Preservation and Long Term Access through NETworked Services)
- CASPER (Cultural, Artistic and Scientific Knowledge for Preservation Access & Retrieval)
- SHAMAN (Sustaining Heritage Access through Multivalent ArchiviNg)
Kahn M. (2010) provides essential guide for planning & preservation of digital information. Gives selected list of organizations concerned with long term retention issues involving digital records & their associated media. It covers

- ARMA International, the Association of Records Managers and Administrators
- Association for Information & Image Management (AIIM’s)
- Association of moving image Archivists (AMIA)
- CAMiLEON
- Council on Library & Information Resources (CLIR)
- Digital Library Federation (DLF)
- Digital Preservation Coalition (DPC)
- Electronic Resources Preservation and Access Network’s (ERPANET’s)
- European Commission on Preservation and Access (ECPA)
- International Research on Permanent Authentic Records in Electronic System
  (InterPARES 1& 2)
- JSTOR; The Scholarly Journal Archive
- National Archives and Records Administration Electronic Records Archives (NARA ERA)
- NDIIPP
- National Information Standards Organisation
 National Partnership for Advanced Computing Infrastructure (NPACI)
 Online computer library centre OCLC
 Open Archival Information System (OAIS)
 PADI Australia
 Records Continuum Research Group (RCRG) Australia
 Research Libraries Group (RLG)
 Society of American Archivists (SAA)
 National Archives U.K.

Gala, B. (2012) in their respective presentation discusses digital preservation; its issues, Digital Preservation Standards & Digital Preservation Management Program and provides list of DPM initiatives at National & International level;

 Digital preservation Coalition (UK)
 National Digital Information Infrastructure and Preservation Programme. (USA)
 Digital Curation Centre (UK)
 Web archiving projects; The Internet archive and PANDORA Project (Australia)
 National Digital Preservation Programme (NDDP)
 National Archive of India
 National Digital Library (India)
Katre D. (2011) provides a comparison between the American and Indian digital preservation programme based on essential building blocks such as National Legislation for Digital Preservation, basic legal framework, national legislation preservation initiatives, digital preservation tools, standard and practiced guidelines, digital repository infrastructure, audit and certification. Selected digital preservation policies of various organisations from UK, Canada and Australia are also analysed. And gives brief idea about preservation initiatives in India:

- National Digital Preservation Programme (NDPP) of India.
  
  Launch by ministry of communications and Information Technology in 2008.

- Centre of Excellence in Digital Preservation, Pune

2.3 Backup:

According to Global Backup Survey (2008-2009) data loss is very common. Data may loss accidentally or intentionally due to natural or manmade disaster (Thrones, 2008). Insurance Agency Inc. (2001) and Ontrack Data International, Inc. (2003) give reasons for losing data are hardware failure, human error, software corruption, computer viruses, theft, hardware destruction etc.

According to Tiwari P. (2012) a good database backup strategy prevents data loss and minimizes system downtime. So library can be able to continue functioning in the trigger event (Smart J., 2012). Data One (2013), Doctor G. (2012) discuss that backup provides protection against all type of disaster.
Schreiber (2010) says in article on ‘The backup bible’ only regular backups protects you from totally losing your important documents. Backup is no simple task; in fact it takes careful planning and strategy.


2.3.1 Backup devices and its issues:

Backuphistory.com provides timeline of evolution of backup device from punch card to blue ray disk /hard disk. It shows when particular backup device or technology began/stopped to be used for backup purpose. Tiwari P. (2012) provides brief outline on various backup devices, importance of backup, redo logs, backup cycle, backup type, strategy etc. Thrones (2008) in his presentation covers backup & its importance, threats to data backup services, backup media etc.

Sahai (2012) provides brief idea about digital storage drives, multimedia drives, NAS etc. Wikipedia (2013) gives detail information about various storage technologies. Driveweb.com (2013) discusses detail evolution of storage devices. Reddy D.G. et al. ( ) describe technology journey from floppy disk to cloud storage with timeline, its advantage & limitation. Yurin M. also describes history of backup in detail from punch card backup to cloud storage with timeline and cases & episode of data lost.

Data tape –LTO 4 & LTO 5 is used for backup (Ultrium.com, 2013). Henriksen et al. (2013) in their DCA project give brief review of various storage media such as
Data tapes, hard disk drives, solid-state drives, optical disks, cloud storage with detail and preservation issues of each. Also provides how to choose it for preservation and recommends some brands. Discuss maintenance issues too.

In hard disk drive, SCSI and SATA technology are most reliable disk drives (Schroeder & Gibson, 2007).

Roadmap of development of LTO tape storage capacity is available on website of ultrium(ultrium.com, 2012). This website provides detail information from first generation (200GB) to eight generation (32GB) LTO tape. Tapes are more suitable for digital preservation. Its life span is higher and error rates are lower (Bradley, 2009).

Ravn-Grove (2010) provides preferred properties of storage media. It covers easy to use, easy to copy, widespread usage, low price and durability. These properties should be kept in mind during planning of backup policy.

DCA (2013) gives points to be consider during investing in storage media such as storage facilities, maintenance, need for expansion, environment issue & budget etc. Hunter (2011) provides guideline on digital storage. It provides points to be consider for ideal storage environment. Pinheiro et al. (2007) give brief idea about preservation issues and tips for different storage media.

Hunter (2011) argues USB and fire wire connected HDDs should not use for archiving and long term storage.

Different surveys show that HDD have an equal risk of failing whether there are constantly in use or switched on & off (Pinheiro, 2007). If HDD found without experiencing any errors in the first year of use, it will remain stable for the next some years. It is known as infant mortality, which means if a HDD is prone to errors and failure it will often happen in the first years of use (Schroeder & Gibson, 2007). It reveals that one should keep a sharp eye on storage and do regular checks up every 3-6 months to find out any errors. According to Wikipedia, HDD are failing due to head crash, bad sectors, circuit failure, motor failure and other mechanical failures etc. Bradley (2009) recommends one should change hard disk after every 3-5 years. Schroeder (2007) says after five years failure rates seem to increase. Henriksen et al. (2013) observe HDDs failing after 3 years of intense use.

Nowadays cloud computing and storage is widely used. With reference to preservation, there is risk of losing data and security of data being leaked or hacked. Hoeven (2012) discusses risk of cloud computing in a preservation context; less control over data, copyright and legal issues, dependency and stability on a specific vendor loss of data due to payment issues, increase in access and availability, upload and migration issues. Findlay (2010) recommends some questions to ask any cloud host company before selection.

Breeding M. (2012) discuss various aspects of cloud computing including cloud storage and provides brief idea about various cloud storage services such as Dura Cloud, Google Docs, Dropbox, Sky Drive, Amazon cloud service - Amazon S3
(Simple Storage Service), Box.net, ADrive etc. Further he argues cloud computing reduces risk and costs but once implemented for major applications result in a higher level of dependency on the provider.

Cloud storage service can handle data on a very large scale. But it costs higher as very large scale storage involves the time and costs of the bandwidth required for a transfer from its local source to a storage provider. Most service providers include charges for incoming & outgoing transfer. Cloud storage provides flexibility, convenient, access and reliability but the costs can skyrocket & logistical tasks to transfer into and out the cloud storage services become more challenging. In some cases, libraries use cloud storage for certain projects, but current pricing models and bandwidth capacities make it a less viable option.

2.3.2 Backup Methods:

There are various backup methods such as Full backup, Incremental backup, Differential backup, Mirror backup, Full PC/Computer backup, Local backup, Offsite backup, Online backup, Remote backup, Cloud backup etc. (www.typesofbackup.com, 2013) This site provides examples and advantages & disadvantages of each backup type. Library should use multiple backup method for baking up their data. Besides it, library must perform full automatic backup regularly.
2.3.3 Backup Process:

Wikipedia (2013) discuss manipulation of data and dataset optimization covering compression, deduplication, duplication, encryption, multiplexing, refactoring and staging etc. It also focuses on backup process and discusses two important points RPO and RTO. Besides this throw light on limitation such as backup window, performance impact, costs of hardware and software, labor and network bandwidth. There should be scheduling, authentication and chain of trust for implementation of backup process. To get desire result of backup – backup validation, reporting, logging, validation and monitoring of backup should be done. In short, key factors should be monitored and historical data should be maintained.

2.3.4 Backup policy:

Kahn M. (2004) argues that backup should be taken regularly otherwise library can loss their valuable data, besides it, recreation of data is very difficult and challenging task. It is time consuming and extremely expensive.

At time of preparing policy document, storage facilities, maintenance, need for expansion, environment issue & budget should be consider. Library should use more than one backup solution to avoid data loss and difficulties in retrieving data. Considering size of data collection, budget, staff expertise & use of data (for access/preservation) etc. ‘backup solution’ should be use/select.

According to DataOne (2013) backup policy should clearly identifies roles, responsibilities, backup media, frequency of backup, file formats etc. And gives
backup’s things to consider at policy making stage. Akhtar et al. (2012) point out that backup plan should cover selection of content, backup type, backup schedule & window, storage of backup and retention policy.


Tendberg Data (2013) in its white paper on backup and archival storage practices provides detail about backup and data protection.

Swineburne University of Technology (2012) gives brief idea about backup policy and procedures.

2.4 Related Studies:

Li, Y. (2011) conducted a National survey on Digital Preservation of Institutional Repository (IR) materials among Association of Research Libraries (ARL) member institutes. Examining the current practices of digital preservation of IR materials, the survey of research libraries reveals the challenges, opportunities of implementing digital preservation for IR in a complex environment with rapid evolving technology, practices and standards. Finding shows that most of research library take actions for digital preservation by developing preservation policies. Awareness regarding preservation also seen in libraries.
Kumar, K. (2014) studies on digital preservation and techniques among Engineering Education Institutional Libraries in Rayalaseema Region on Andhra Pradesh. A study shows that hard disk is widely used for preservation. Majority of librarians prefer refreshing & emulation as preservation techniques. HTML & PDF formats are used for digital content. Most of libraries face IPR issue in digital preservation.

Zaveri, P. (2012) has done research on disaster management in libraries in India. She covers all types of disasters including technological. And focuses on back-up and preservation of digital content.

Kabooza conducted global back-up survey covering 4257 respondents from 129 countries during 2008-09. It reveals that people neglect backup needs. Most of respondents do not take regular backups. At the same time 66% have had critical data loss. This is alarming and shocking since the risk of losing important digital content, due to virus, hard drive crashes, physical damage or theft is so high.

Lundgren, M. (2012) studies on data backup practices in an academic environment. Study reveals that nowadays external hard disk and online backup-drop box are used for backup. None of participants used CD, DVD and Blue ray to backup their data. Besides this university server were not used. It has observed that most backup were not preserved in any way. Respondent considered backup to be of importance. But they were not taking backup regularly. They take backup of data only whenever they felt worried for some particular file & folder.
Mugoh Leon et al. (2011) conduct survey on continues data protection architecture as strategy for Reduce Data Recovery Time. According to them backup on magnetic tape requires high data recovery time. CDP-Disk based backup solution ensures faster data retrieval. Some backup vendors are bundling CDP in to their backup software. CDP continuously backup data in real time.
References:


Doctor, G. (2012). *Disaster management*. In refresher course in LIS. V.V.Nagar; S.P.University. (Presentation)


Gala, B. (2012). *Digital library preservation & archiving*. In refresher course in LIS. V.V.Nagar; S.P.University. (Presentation)


Chapter 2  Review of Literature


Smart, J. (2012). *Backup*. In refresher course in LIS .V.V.Nagar; S.P.University. (Presentation)


Tiwari. (2012). *Basics of data storage system* (Presentation). In refresher course in LIS. V.V.Nagar; S.P.University.


Watson. (2009) *Preservation concerns in the e-resource environment*. In Managing the Transition from Print to Electronic Journals and Resources: A Guide for Library and Information Professionals (eBook) (Downloaded from INFLIBNET site)

