Preservation of Knowledge in the Electronic Age: An overview of University Libraries in North East India

Senthony Narzary¹, Rudra Singh Daimari²

¹Assistant Librarian, APJ Abdul Kalam Central Library, bengtol College, Bengtol, Chirang, Assam
²Librarian, APJ Abdul Kalam Central Library, Bengtol College, Bengtol, Chirang, Assam

Abstract
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. This paper discusses the importance and need of preservation of information resources in libraries, factors that affect degradation and deterioration of library materials and methods of tackling them. It also discusses the preservation problems that are associated with the digital information and the challenges faced by the university librarians of North East India in preserving the intellectual content contained in the digital media.

Keywords: Preservation; Digital Preservation; Digital information; Non-print Material; Preservation Policies; North East India University Libraries

1. Introduction
Libraries have always struggled against the physical destruction of their collections. Fires, floods, earthquakes, and wars have damaged the holdings of countless libraries, destroying forever much of the recorded history of human civilization. But library materials also fall victim to slow decay caused by acid content in paper, insect infestation, improper storage or handling, and excessive heat, mildew, humidity, and air pollution. The slow decomposition of library materials is a universal problem. To ensure that library materials remain available to present and future generations of library users, libraries engage in a variety of preservation efforts. These efforts include the conservation of original materials and the transfer of information from original materials to more durable formats. Preservation does not simply happen on its own, a well thought out plan must be drawn and managed. According to fifth Law of Library Science “Library is a growing organism” Libraries acquire materials of all kinds continuously, and promote the use of these acquired materials. Hence more and more number of users wants to access these materials. As more and more number of users use these materials, they are more likely to be damaged. To prevent this deterioration of materials which may affect the further retrieval of the contents, we need to adopt an array of appropriate management strategies. Environmental conditions and methods of storage have a great influence on the preservation of documents. Control of the
environmental conditions and the provision of good storage conditions constitute the best preventive measures.

Today, many libraries and archives are in the process of ‘going digital’. Moreover, as of today no one knows how to ensure the long-term preservation of multimedia documents nor how to ensure the integrity of documents that may have many links to other documents that may be anywhere in the world. Yet the fact remains that several key issues concerning the long term preservation of digital technologies remain unsolved. The digital technology is well known and its adoption by libraries and archives seems inevitable, inexorable and well-motivated. But problems are the fragility of digital media even more intractable, is the rate at which computer hardware and software become obsolete. Many cases have been cited in which valuable data has already been lost because of obsolescence. While backup, system redundancy, and byte replication may be used by delivery organizations and digital preservation organizations, these actions alone are not sufficient for digital preservation. These problems have been exercising the library and archive communities for some time but as yet no one solution or set of solutions has been reached. Therefore, solutions need to be found urgently if not to sink in what Rothenberg calls ‘technological quicksand’.

2. Preservation
Preservation is a fundamental recordkeeping issue. Preserving records goes well beyond the conventional view of repairing damage after it happens. It focuses on preventing damage and keeping records useable for as long as we need them. This is done by examining the factors that cause deterioration.

3. Causes of Deterioration
Causes of deterioration may be external or internal. External causes are physical, chemical, biological, improper storage, unauthorized exposure, risky handling, theft, natural calamities. Internal causes are: poor quality of paper or materials used, the chemicals used for printing and binding. In this context the external factors and the measures to counter the effects are discussed here.

In this context the external factors and the measures to counter the effects are discussed here.

3.1. Physical Agents
- Temperature, Relative Humidity, Light, Sound (Vibration), Magnetism.

Temperature is important especially in a tropical country like India, where the temperature fluctuates from one extreme to another during the year and even during a single day, shows at times extreme variations. These drastic fluctuations in temperature forces the documents adopt themselves to the changing conditions too frequently, and the strain of these variations will have an impact on the strength of the materials. In a humid atmosphere, the danger to library materials comes from microorganisms which thrive best under such conditions. Moreover paper is a hygroscopic material and absorbs moisture from a moist air. Such moisture can weaken the strength of paper documents, spreading of inks used for printing. If high humidity is combined with bad ventilation and improper lighting, the resulting harm to library materials is indescribable.

In a tropical country like India, the abundance sunlight poses a great problem. Because sunlight contains quite a high proportion of ultra-violet radiation which is so powerful as to induce degenerative chemical changes in organic material on which it is incident. Paper is mainly made up of cellulose which
constitutes long chains of molecules made of carbon, hydrogen, and oxygen. Ultra-violet radiation breaks up the links in the long molecular chains and thus brings about structural decay of paper. The strength of paper decreases and it gets discoloured. Inks fade and printing becomes faint. The leather of bookbinding tends to get hard and brittle. Ultra-violet rays of wavelength shorter than 300 millimicrons are particularly detrimental to the cellulose of paper. It is desirable to prevent sunlight from falling directly on books and other library materials.

3.1.1. Physical Agents and their Preventive Measures
- Maintain constant temperature (22 –23°C) and Relative Humidity (45 – 50%)
- Proper ventilation should be maintained
- Air-conditioning is the best solution
- To control excess humidity install dehumidifiers, or placement of silica gel in appropriate places.
- Use sun filters, lead bulbs, acetate foils etc. to control the powerful radiant light energy
- The levels of light (should not exceed 50 Lux ) should be decided based on the stack area and the user area
- Handle carefully the documents in the library while transporting from one floor to another
- Avoid using mobile storage system which may create friction and damage documents
- Libraries should be located away from subways, trains, automobiles, etc. to avoid the effect of vibration from outside sources
- Stray magnetic fields are the natural enemy of magnetically recorded information. Avoid using magnets, magnetic notice boards, etc.

3.2. Chemical Agents
- Dust and Dirt
- Internal Acidity of Paper and Ink
- Air Pollution and Atmospheric Gases
   The environment in industrial cities is highly polluted. It contains pollutants like dust, smoke, coke dust, fly ash, salt particles, calcium, ammonium sulphate, nitrates, chlorides, solid oxides, soot tars, and gases like carbon monoxide, nitric oxide, nitrous oxide, sulphur dioxide, ozone, olefins, aromatic hydrocarbons, aldehydes, ketones, paraffins, hydrogen sulphide, halogen compounds and ammonia. All these factors lead to increase the acidity of the documents. Deacidification methods can be used to remove the acid content and increase the longevity of documents. The Library of Congress estimates that deacidification can prolong the life span of paper-based library materials by 250 to 300 years.

3.3. Biological Agents
   Biological agents thrive on the organic matter they find in library materials. Absence of proper ventilation, darkness, high temperature and relative humidity encourage their spread. The biological agents can be grouped into macro organisms and micro organisms.

3.3.1. Macro organisms
- Silver fish, Book lice, Book worm, Cockroaches, White ants (Termites), Rodents, Man
   There are many pesticides like (DDT; Pyrethrum; Para dichloro benzene etc.) available in the Market
These pesticides may be sprayed periodically. There are agencies specialized in this area that may be consulted in case one has any specific problems. The control of termites requires specialised attention in libraries. As termite infestation starts from the soil, creation of chemical barrier around the building using crude creosote in kerosene or Dieldrex in water is recommended. Wooden structures are to be coated with creosote oil and solignum. Constant vigil against termite infestation is to be maintained. Care should be taken that books and other documents on the shelves are not directly sprayed upon. Problem of rodents can be controlled by trapping, stomach poisoning (zinc phosphide and arsenic oxide) and sealing of crevices, cracks, and entry points. Markings on the documents by pencil or pen by users and using the book as a pillow for protecting their heads during monsoon also contributes to the deterioration process of documents. It is a very difficult and delicate issue. This can be prevented by educating the users and displaying notices explaining the importance of books at prominent places in the library.

3.3.2. Micro organisms

• Fungus, Mildew and Mould

A fungus attack can be easily identified by fuzzy growths, mostly coloured, which seem to spread out from a point. About 100 different species of fungi have been found one time or another to infest organic materials in libraries. Some important fungi which attack library material, especially paper are Pencillium, Fusarium, Trichoderma, Alternaria, Citromyces, Aspergillus and Monilia. The proper identification of fungi is done under microscope. There are fungicides like Mercuric chloride in spirit or alcohol or in water; Para nitro phenol 3% in water; Thymol + Sodium chloride; Sodium silicate + Sodium penta chloride; Chloro phenol; Hydroxy quiniline, etc. may be sprayed periodically. Care should be taken that books and other documents on the shelves are not directly sprayed upon. Fumigation methods (vacume and ordinary) are the most effective ways of sterilising of documents from fungus, mould and mildew. In vacume fumigation method, following chemicals are used.

• Ethylene oxide + CO2 1:9 by weight for 10m3 i.e. about 4.5 kg for four hour
• Ethylene chloride + CO2
• Methyle formate + CO2
• Carbon tetra chloride
• Hyadrocynic gas
• Formaldehyde (250 gms/cubic area) at 30 0°C constant temperature for 48 hrs.
• Killopetra (Ethylene dichloride + Carbon tetra chloride (3:1) i.e. ½ litre /2m3 at 23.850°C for 36 hrs.
• Para dichloro benzene, 1 kg = 1m3 for 72 hrs.
• In ordinary fumigation method, following chemicals are used.
• Thymol (120 gms = 1m2 for seven days)
• Fomaldehyde

3.4. Materials Handling

• Storage system, Exhibition, Transportation, Photocopying

Libraries should have proper storage system depending upon the type of documents such as books, periodicals, microfilms, micrifiches, CD-ROMs etc. and size of documents such as maps, atlases, globes, etc. When documents are removed for exhibition, care should be taken to avoid direct exposure to ultraviolet radiation and minimise the transportation. Indiscriminate photocopying should be avoided.
3.5. Theft and Vandalism
- Tag materials with magnetic strips. These strips will trigger alarms if users try to carry the materials through electronic gates at library exits without properly checking out the items at the circulation desk.
- Install closed circuit TV to monitor the users' movements.
- Use wire mesh to cover windows.
- Deploy security guards at the library entrance.

3.6. Fire
- Install fire or smoke alarm system.
- Keep fire fighting equipments in designated places in the library.
- Provide fire fighting training to all employees in the library.

3.7. Natural Disasters
- Flood, Cyclone, Earthquakes
  It is very difficult to predict the natural disasters. There should be a contingency plan which should form part of the preservation policy aimed to minimize the loss and rescue the materials to safer places when disasters strike. For that it is vital to know which procedures to follow, to test them and include them in periodic security exercises involving all of the library staff. An up-to-date address list must be kept indicating whom to call in case of disaster (vital installations such as security, hospital, fire fighting unit, etc.).

4. Digital preservation
It is defined as the managed activities necessary: 1) For the long term maintenance of a byte stream (including metadata) sufficient to reproduce a suitable facsimile of the original document and 2) For the continued accessibility of the document contents through time and changing technology.

4.1. Audio and Visual Materials
Not only paper-based materials risk deterioration on library shelves. Similar dangers confront audio and visual library materials, such as sound recordings, photographs, films, and videotapes. For example, nitrate-based film stock was the only available format for motion-picture production until 1951, but the nitrate in this type of film causes it to decay very quickly, even in controlled settings. Many have been lost or destroyed, but a vast number have simply decomposed beyond repair. Libraries and archives preserve nitrate-based films by transferring the images to a more resilient, acetate-based film stock. They preserve other audio and visual materials in similar ways. For example, original sound recordings are preserved by transferring them from delicate and unstable wax cylinders or magnetic tapes to newer digital formats such as CD-ROMs. In addition to preserving their materials from deterioration, libraries must guard against the obsolescence of machine-readable materials—materials that are read and interpreted by machines. Many valuable documents in machine readable materials were first recorded in formats that have now become obsolete. Machines able to play back the recordings either no longer exist or are so rare that they are not practical for use in libraries or even for storage in archives.
4.2. Computer Data

Computer software and hardware pose an additional complex problems and challenge to the preservation efforts of libraries. Because common standards for computer software and hardware change so quickly, vast amounts of information stored in obsolete computers can no longer be accessed using modern equipment. As a result, libraries risk forever losing access to valuable computer documents such as government statistical data and geological surveys. To ensure that original computer data remain accessible using contemporary equipment, libraries and archives must continually transfer these data to new formats, which is extremely costly and time-consuming process. Most library conservators and archivists can transfer and preserve only those materials that they determine are of enduring value. As the quantity of computer-based records increases each year, the task of identifying which electronic materials warrant preservation becomes increasingly difficult. The major challenge before the librarians is how to archive the online versions of print journals and one has to work out a strategy as how to make this possible. The digital media like computers, hardware, software, floppies, CDROMs, databases etc. are affected not only by environmental factors and biological viruses but also by various types of computer viruses (programs) that affect and corrupt systems, files, floppies, CDs etc. New problems like cyber crimes, hacking etc. have become order of the day. Therefore, librarians have to be more vigilant by evolving appropriate, effective and efficient tools like firewalls, passwords, periodical backups etc. to protect the information available in network and non network environments.

4.3. Components of Digital preservation

The following components are necessary to achieve digital preservation:

• **An independent organization with a mission to carry out preservation**

As noted in a recent CLIR survey, *E-Journal Archiving Metes and Bounds: A Survey of the Landscape*, the first indicator of an archiving program’s reliability is that it ‘have both an explicit mission and the necessary mandate to perform long term e-journal archiving’. The mission creates an environment conducive to the specialized planning and infrastructure needed to support digital preservation.

• **A sustainable economic model that can support preservation activities over the targeted timeframe**

The actual costs of long-term digital preservation are difficult to determine with accuracy at this early date. A preservation organization may choose to preserve all content in its archive for the same amount of time, or may preserve different items for different lengths of time. The retention time of every item preserved should be set at the point the content is acquired and imported into the archive. The preservation organization must have an economic model that provides enough funds to move the content into the archive, continue its ongoing technology watch, and implement intermittent preservation activities on the preserved content for the stated retention time period.

• **Clear legal rights to preserve the content and relationships with the content providers**

A digital preservation organization must have legal rights to preserve the digital content ingested into its archive. These rights must be obtained in advance of the acquisition of the content and must delineate the scenarios under which the preserved content is to be made accessible. Because these legal agreements must always remain with the archived content, they should be preserved in the archive along with the digital content. Two of the key elements of digital preservation are that the content must remain usable and accessible over the long term. In order to fulfil these requirements, the preservation
organization must understand the structure and format of the content. For example, while staff at the preservation organization need not understand physics to preserve physics journals, the staff must understand how scholarly journals – and physics journals in particular – are typically structured, and how a particular journal is structured. With this knowledge, the preservation organization can develop the tools necessary to automatically read the source files that comprise the journal and to retrieve the bibliographic data – the data which ensures that the content is discoverable in the future.

- **Relationships with the eventual users of the content**
  As noted in the OAIS standard, an archive exists to support a designated community, ‘an identified group of potential consumers who should be able to understand a particular set of information. The designated community may be composed of multiple user communities. The preservation organization must have an ongoing relationship with its designated community and must have a way of eventually delivering usable, authentic, and discoverable content to that community.

- **A preservation strategy consistent with best practices and a technological infrastructure able to support the selected preservation strategy:**
  Best practices and standards in digital preservation are continuing to evolve, but a number of guidelines do exist, including:
  - OAIS (Reference Model for an Open Archival Information System, ISO 14721: 2003) a high-level framework for designing a preservation organization.
  - PREMIS (Preservation Metadata: Implementation Strategies) – a data dictionary and documentation describing the metadata necessary for preserved digital content.
  - TRAC (Trustworthy Repositories Audit & Certification: Criteria and Checklist) – a set of digital preservation best practice criteria that can be used to evaluate repositories
  - DRAMBORA (Digital Repository Audit Method Based on Risk Assessment) – a risk management methodology that allows a repository to run an internal audit in order to assess its capabilities, weaknesses, and strengths.
  - Nestor (Network of Expertise in Long-Term Storage of Digital Resources) Catalogue of Criteria for Trusted Digital Repositories—a checklist similar to TRAC to assess the technical and organizational trustworthiness of a digital repository.
  - DPC Handbook (Digital Preservation Coalition) – a detailed guide to the management of and long-term access to digital objects

- **Transparency about preservation services and strategies, clients, and content**
  Just as the OAIS standard details how an organization should define its designated community and its clients, the CLIR report details a set of indicators of reliable digital preservation repositories. The report notes that a preservation organization must ‘be explicit about which scholarly publications it is archiving and for whom offer a minimal set of well-defined archiving services’. A preservation organization should clearly communicate its preservation methodology to its designated community and provide the community with a method of auditing the preserved content.
4.4. Near-term access protection and long-term content preservation

The steps involved in both near-term access protection and long-term content protection through preservation can be placed along a continuum. The key goals of digital preservation include:

- **Backup**
  
  Backup has long been understood to be a requirement for protection of near-term data access. It is imperative for business continuity and it is necessary to ensure that access to content in the near term will not be interrupted for any length of time. Backup is typically implemented with commercial software that allows users to retrieve files backed up at specific points in time. If special software or hardware is required to access the content and if it has been compressed via a proprietary technology, the long-term future accessibility and authenticity of the content cannot be assured.

- **Access system redundancy**
  
  Access system redundancy is an excellent way to ensure that there is little interruption to near-term, ongoing access but it does not alone guarantee usability, authenticity or accessibility of the content over the long term as technology and data format evolves.

- **Byte replication**
  
  Byte replication is a process whereby identical, multiple copies of files, file systems, or websites are created. They may be written to other online computers or to offline media. These replicas are typically held in diverse geographic locations and specialized software is not needed to access the content. This diversity in copies and location, together with the lack of reliance on software, ensures that byte replicas should provide content that is authentic and usable for as long as the file formats remain readable. However, simple byte replication includes no provision for ensuring the content is usable when the file formats are no longer current, nor is there any inherent provision for ensuring that the content remains discoverable.

5. Methods of digital preservation

Migration and emulation are the two primary strategies used for long-term preservation.

- Migration involves transforming digital content from its existing format to a different format that is usable and accessible on the technology in current use. Migration is a strategy that requires a deep understanding of the content being preserved.
- Emulation involves developing software that imitates earlier hardware and software. Emulation is a more technology-based strategy, requiring a deep understanding of existing hardware and software.

5.1. The Technical Issues

Each of three elements of digital technology: Media, Hardware and Programming

5.1.1. **Media**: Archivists and librarians, responsible for maintaining a continuous record of human activity and thought for present and future generations, view the life expectancy of storage media in terms of hundreds of years. Archival-quality silver microfilm can last up to 200 years. Yet, some magnetic tape can become unreliable for archival storage after only five years. Likewise, many optical disk media of average quality, including CD-ROM’s, are not reliable after five years. Optical and magnetic media are currently the two most common repositories of digital information.
• **Optical disks**, unless stored properly, are susceptible to a potential breakdown between their reflective backing where data are stored and their protective, transparent surface. Also, with improper storage, the transparent surface of optical disks could become cloudy, hampering the ability of machinery to read the disks. Surface scratches are another hindrance to the retrieval of data from optical disks. Other types of optical media, such as WORM (Write once, ready many times), are more reliable than CD-ROM's and guaranteed for up to 100 years by some manufacturers.

• **Magnetic tapes** can become brittle, causing magnetic coating to separate from its backing. Also, magnetic tapes are susceptible to interference from magnetic forces in the environment that may cause errors and omissions in recorded data. Studies by the National Media Laboratory have determined that average-quality magnetic tape, kept at a constant room temperature, becomes unreliable as a storage medium in five years or less. The National Institute of Standards and Technology (NIST) estimates "the longevity of modern magnetic tape to be about 20 years under ideal storage conditions."

### 5.1.2 Hardware:
If a digital file is recorded on an optical or magnetic disk and there are no disk drives to read that particular size and format, the information on that disk is lost.

### 5.1.3. Programming:
Software advances are occurring as fast as hardware advances, bringing rapid obsolescence to countless programs and languages. Unless documentation for out-of-date software is maintained into the future, stored data become as unreadable.

### 6. What Can We Do?
One of the most important steps that can be taken is to give careful consideration to the long-term preservation needs of digital information when it is first created. Librarians and archivists are urging those who generate data to set priorities and choose the portions that deserve permanent preservation. But choosing is only the first step. In the past, librarians and archivists carefully selected the paper-based documents that seemed likely to have long-term value. That relatively small portion of materials received appropriate preservation treatment, usually after careful assessment of the condition of the paper. Selection of digital materials for long-term preservation is much more difficult, because obsolete hardware and software can render such materials unusable, even though the material itself is in good condition.

### 7. University Libraries of North East India
Although there are more than 50 (fifty) Universities comprises of Central, State, Deemed, Private and National Institutes of Importance, it will still take time for libraries in the North East India to reach a state where a Digital preservation in the libraries is fully implemented. Despite of the factors impeding the consortia initiatives, university libraries in the North East India are coming up to be at par with the mainstream. But it is clear that all the libraries are in the initial stages of automation and creating complete database of all records is also not complete in most of the libraries. Some of the newer universities have not even started whereas some have completed creating database of their records. Considering the geographical topography and lack of infrastructure the North Eastern states are a unique cases with embedded complications for suitable connectivity. However, most libraries do not have
specific software and hardware for Digital preservation. Another important aspect is the numbers of staff are not trained in handling Digital Library materials and information technology. For a fully Digital preservation of the library material most of the staff should have some Knowledge of IT and Digital Preservation Techniques.

8. Conclusion

The scholarly community, like the general public, is moving rapidly into this digital environment and the digital scholarly content being created today must be preserved for the future – without guaranteed access to this record over the very long term, future research will be hampered. Traditional preservation responsibilities and methodologies are not applicable to digital content, where physical copies are not delivered to libraries. Instead, in order to meet the unique preservation needs of digital content, libraries and publishers, two key participants in the scholarly communications environment, must join together to ‘invest in a qualified archiving solution’.

Preservation aspects are given a least priority in many libraries. The modern librarian is faced with a complex problem as to deal not only with the traditional media but also with modern media like video recordings, photographs, microfiches, microfilms, CD – ROMs, computer software, online databases and their quick obsolescence of associated technologies. It is the duty of the librarian to preserve and conserve the intellectual content of the documents for the generations to come irrespective of media and technological changes.

With the growth of digital technology, the library and archival community faces wonderful opportunities to provide new forms of access to an increasing amount of information, and formidable challenges in managing that information, especially in ensuring the long-term availability of and ready access to data in digital formats. Traditional approaches to preservation cannot address the whole range of issues that arise from data created in digital formats, carried on a variety of media, such as magnetic tape, CD-ROMs, and computer hard drives, and retrieved through a large number of software programs that routinely and rapidly become obsolete.

References:


