

*Full Length Research Paper*

# Digital preservations in state universities of Andhra Pradesh: A survey

T. Raghunadha Reddy\* and K. Kumar

Sri Venkateswara College of Engineering and Technology (SVCET), Chittoor-517001, India.

Accepted 17 December, 2012

**The purpose of this paper is to find out the awareness of librarians about digital preservation. This article intends to describe the emerging role and fit of digital types of information in scholarly practice and to observe the impact of metadata standards on their learning. A structured online questionnaire was arranged to gather data from the librarians in state universities. Based on the findings of the survey, it is concluded that there exists adequate awareness of digital preservation among librarians. Document scanning program seems to be the major available source system in state universities of Andhra Pradesh. Dublin core metadata are used in most of the state universities' libraries.**

**Key words:** Digital preservation, universities, librarians.

## INTRODUCTION

The direction, digital preservation is known (Robert Neumayer, 2011), but the details and complex relationships that must be resolved are still being pursued. It is on record that the arrays of resources and expertise are making inroads into the solution of all this digital complexity. Almost all of the digital preservation exists within established institutions and have dedicated staff with technological expertise up to the task of wrestling with aspects of these issues. These include, for example, national libraries, national archives, institutional repositories, universities and other places of learning and research, media, cultural museums and archives. In other words, they are to develop a sustainable standard digital preservation management and storage system for a collection institution that happens not to be one of the world leaders in digital preservation. The digital libraries in state universities of Andhra Pradesh were established according to the state's tradition with the aim of locating, collecting, preserving, digitalizing and making accessible the collected knowledge of Andhra Pradesh State.

### Digital preservation

Digital preservation refers to the series of managed activities necessary to ensure continued access to digital

materials as long as necessary. Digital preservation (<http://www.dpconline.org/advice/preservationhandbook/introduction/definitions>) is defined very broadly for the purpose of this study and refers to all of the actions required to maintain access to digital materials beyond the limits of media failure or technological change. These materials may be records created during the day-to-day business of an organization; "born-digital" materials created for a specific purpose (example, teaching resources); or the products of digitization projects. This excludes the potential use of digital technology to preserve the original artifacts through digitization. See digitization definition below:

1. Long-term preservation - Continued access to digital materials, or at least to the information contained in them, indefinitely.
2. Medium-term preservation - Continued access to digital materials beyond changes in technology for a defined period of time but not indefinitely.
3. Short-term preservation - Access to digital materials either for a defined period of time while use is predicted but does not extend beyond the foreseeable future and/or until it becomes inaccessible because of changes in technology.

## REVIEW OF LITERATURE

Financial sustainability is an important attribute of a

\*Corresponding author. Email: [raghuthima@gmail.com](mailto:raghuthima@gmail.com).

trusted, reliable digital repository (Palaiologk, 2012). The authors of this paper use the case study approach to develop an activity-based costing (ABC) model. This is used for estimating the costs of preserving digital research data and identifying options for improving and sustaining relevant activities. The model is designed in the environment of the Data Archiving and Networked Services (DANS) institute, a well-known trusted repository. The DANS-ABC model has been tested on empirical cost data from activities performed by 51 employees in frames of over 40 different national and international projects. Costs of resources are being assigned to cost objects through activities and cost drivers. The 'euros per dataset' unit of costs measurement is introduced to analyze the outputs of the model. Fenders, managers and other decision-making stake-holders are being provided with understandable information connected to the strategic goals of the organization. The latter is being achieved by linking the DANS-ABC model to another widely used managerial tool-the Balanced Scorecard (BSC). The DANS-ABC model supports costing of services provided by a data archive, while the combination of the DANS-ABC with a BSC identifies areas in the digital preservation process where efficiency improvements are possible.

This article looks at the digital preservation in the Punjab Region of India and Pakistan through the creation of the Punjab Digital Library (PDL), which began operations in 2003 (Gurpreet Singh, 2012). The digital library was established to preserve the region's heritage with the aim of locating, collecting, preserving, digitizing, and making accessible the accumulated wisdom of the Punjab Region. The technical challenges encountered in building up the collection in the library are explored.

### Objectives

1. To find out the awareness of librarians about digital preservation.
2. To find out the frequency of using digital data (prepare tables on the opinion of the users)
3. To describe the emerging role and fit of digital information types in scholarly practice
4. To understand the impact of budget on scholarly practice
5. To find out the impact of metadata standards on their learning
6. To find out the hindrances and problems faced by the users while accessing and using digital preservation

### Limitations

1. The study covers only 11 state universities in Andhra Pradesh.
2. The study is related to the responses received through

online questionnaire.

### METHODOLOGY

The present study was carried out to assess the knowledge of digital preservation among the librarians of the state universities of Andhra Pradesh ([http://www.ugc.ac.in/inside/State\\_University\\_November2011.pdf](http://www.ugc.ac.in/inside/State_University_November2011.pdf)). Population sampling method was used in this study and structured questionnaire was mailed to all the e-mail IDs of the respondents, collected from the respective library websites. 32 questionnaires were mailed to librarians, out of which 11 responses were received. These constitute 34.37% of the total response and same was used for analysis purpose, mentioned in Table 1. Figure 1 shows the various types of universities used in this study.

### DATA ANALYSIS

The data collected from the questionnaire are analyzed to fulfill the stated objectives. Statistical analysis techniques such as frequency distribution, percentages are used.

#### Awareness of digital preservation

Table 2 reveals that 11(100%) of the respondents are aware of digital preservation.

#### Digital data

Table 3 presents the digital data status in the universities; majority of the libraries are having below 1 terabyte (45.45%), followed by between 1 to 50 terabyte (18.18%); only 9.09% is having between 101-150 terabyte, 151-200 terabyte and above 200 terabyte.

#### Digital information types

Table 4 explains the availability of the digital information types in the universities. Majority (90.91%) are having document, e-books, e-journals and CD/DVD, followed by images (81.82%); only 9.09% has scientific data.

#### Source systems

Table 5 presents the availability of the source systems in the universities. Majority (27.27%) have document scanning programs, followed by file systems, Internet and digitization programs (18.18%) equally; only 9.09% has electronic document management systems and computer assisted design equally.

#### Preservation policies

Table 6 reveals the preservation policies in the

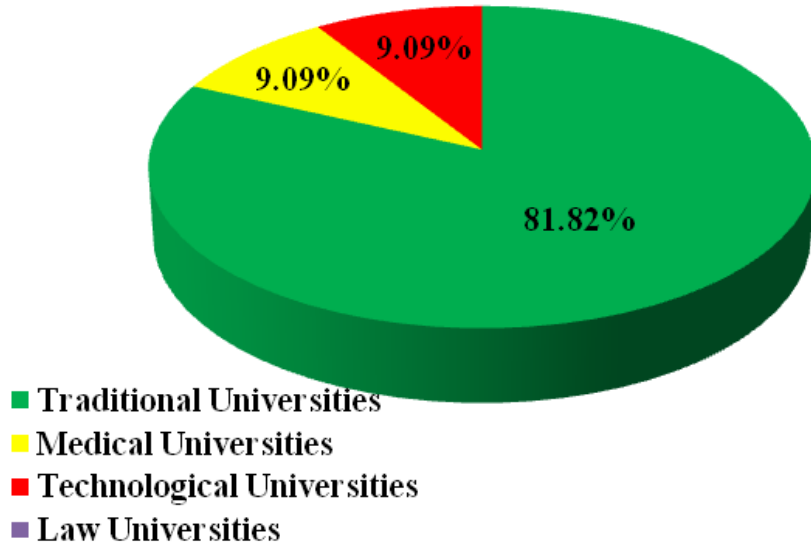


Figure 1. University-wise questionnaire received.

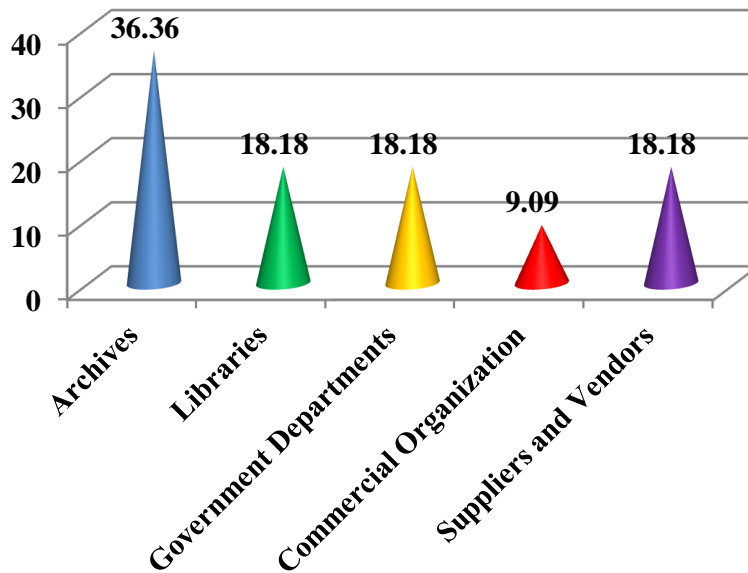


Figure 2. Preservation policies.

universities. Majority (36.36%) are having archives, followed by libraries, government departments, suppliers and vendors (18.18%) equally and only 9.09% has commercial organization. The pictorial representation is given in Figure 2.

**Separate budget for digital preservation**

Table 7 explains the opinion on the allocation of budget for digital preservation in the universities. Majority

(63.64%) said it was sufficient and only 36.36% said it was insufficient.

**Essentials of digital collection**

Table 8 reveals the opinion on the essentials of digital collection in the universities. Majority (63.64%) indicated that records have not been scratched and only (27.27%) indicated that records are accessible for up 25 years and are integrated into content delivery systems.

**Table 1.** Response from the librarians.

S.No.	State universities	Distributed	%	Received	%
1	Traditional Universities	21	65.6	9	81.82
2	Medical Universities	3	9.38	1	9.09
3	Technological Universities	5	15.6	1	9.09
4	Law Universities	2	6.25	0	0
Total		32	100	11	100

Response Rate: 34.37%.

**Table 2.** Awareness of digital preservation.

Sl.No.	Awareness of digital preservation	No. of respondents	Percentage
1.	Yes	11	100
2.	No	0	0
Total		11	100

**Table 3.** Digital data status in the universities.

S. No.	Digital data	No. of respondents	Percentage
1	Below 1 terabyte	5	45.45
2	Between 1 to 50 terabyte	2	18.18
3	Between 51-100 terabyte	1	9.09
4	Between 101-150 terabyte	1	9.09
5	Between 151-200 terabyte	1	9.09
6	Above 200 terabyte	1	9.09
Total		11	100

**Table 4.** Digital information types.

S.No.	Digital information types	No. of respondents	Percentage
1	Documents	10	90.91
2	Images	9	81.82
3	Databases	4	36.36
4	Audio	5	45.45
5	Websites	7	63.64
6	Video	3	27.27
7	Electronic mails	7	63.64
8	Scientific data	1	9.09
9	Software	7	63.64
10	Electronic books	10	90.91
11	Electronic journals	10	90.91
12	CD/DVD	10	90.91

N = 11.

### Using metadata standards

Table 9 presents that majority (63.64%) are using the Dublin core metadata and only 18.18% use METS and MARC equally. Other Standards are not used.

### Constraints of digital preservation

Table 10 describes that majority (63.64%) are having backup, authentication equally for the constraints in digital preservation followed by duplication (54.55%) and

**Table 5.** Availability of source systems.

S.No.	Source Systems	No. of respondents	Percentage
1	File systems	2	18.18
2	Document scanning programs	3	27.27
3	Internet	2	18.18
4	Electronic document management systems	1	9.09
5	Digitization programs	2	18.18
6	Computer assisted design	1	9.09
Total		11	100

**Table 6.** Preservation policies.

S.No.	Preservation policies	No. of respondents	Percentage
1	Archives	4	36.36
2	Libraries	2	18.18
3	Government departments	2	18.18
4	Commercial organization	1	9.09
5	Suppliers and vendors	2	18.18
Total		11	100

**Table 7.** Budget for digital preservation.

S.No.	Budget	No. of respondents	Percentage
1	Sufficient	7	63.64
2	Insufficient	4	36.36
Total		11	100

**Table 8.** Essentials of digital collection.

S.No.	Essential of digital collection	No. of respondents	Percentage
1	Maintains authenticity	6	54.55
2	Check records have not been damaged	7	63.64
3	Digital information management	4	36.36
4	Content to deal with technical obsolescence	4	36.36
5	Ensures records are accessible for up 25 years	3	27.27
6	Is able to store many different types of content	4	36.36
7	Handles a wide variety of file formats	7	63.64
8	metadata standards	5	45.45
9	Retrieves content using full text	6	54.55
10	Check for duplicate items	4	36.36
11	Integrates with content delivery systems	3	27.27
12	Support emulation to deal with technical obsolescence	5	45.45

N = 11.

Internet (45.45%); only 27.27% is having power and legal issues for constraints.

digital information in India.

### List of Digital Preservation

Table 11 presents some of the institutions preserving

### Digital preservation soft wares

Table 12 presents the major digital preservation Soft

**Table 9.** Using metadata standards.

S.No.	Metadata standards	No. of respondents	Percentage
1	Dublin Core	7	63.64
2	METS	2	18.18
3	MARC	2	18.18
4	EAD	0	0.00
5	PREMIS	0	0.00
6	MODS	0	0.00
7	TEI	0	0.00
8	DDI	0	0.00

N = 11.

**Table 10.** Constraints of digital preservation.

Sl.No.	Constraint	No. of respondents	Percentage
1	Backup	7	63.64
2	Power	3	27.27
3	Duplication	6	54.55
4	Authentication	7	63.64
5	Internet	5	45.45
6	Legal Issues	3	27.27

N = 11.

**Table 11.** List of digital preservation.

S.No	Institutions	Websites
1	Digital Archive of National Institute of Technology Rourkela	<a href="http://dspace.nitrkl.ac.in/dspace/">http://dspace.nitrkl.ac.in/dspace/</a>
2	Electronic Theses and Dissertations of Indian Institute of Science	<a href="http://etd.ncsi.iisc.ernet.in">http://etd.ncsi.iisc.ernet.in</a>
3	Open Access Repository of IISc Research Publications	<a href="http://eprints.iisc.ernet.in/">http://eprints.iisc.ernet.in/</a>
4	IDRC Digital Library	<a href="http://idl-bnc.idrc.ca/">http://idl-bnc.idrc.ca/</a>
5	Digital Repository of IIT Bombay	<a href="http://dspace.library.iitb.ac.in/dspace/">http://dspace.library.iitb.ac.in/dspace/</a>
6	DSpace at National Centre for Radio Astrophysics	<a href="http://ncralib.ncra.tifr.res.in:8080/dspace">http://ncralib.ncra.tifr.res.in:8080/dspace</a>
7	Raman Research Institute Digital Repository	<a href="http://dspace.rii.res.in:8080/dspace/">http://dspace.rii.res.in:8080/dspace/</a>

wares available (<http://www.greenstone.org/software> details; <http://www.eprints.org/software> details; <http://www.dspace.org/software> details; <http://fedora-commons.org/software> details; <http://koha-community.org/software> details).

## Conclusion

Digital preservation is achievable in a country where policy frameworks, institutional frameworks, information infrastructure, trained manpower and financial resources are adequately available. Such systems are likely to be

componentized. Market and technology trends analysis (2007), mix-and-match solutions. They will need to be scalable, particularly to handle the predicted large volumes of content, and also to handle high ingest rates. In addition, they will need to handle a wide range of formats from a variety of sources, and preserve the information contained therein for up to 50 years. For universities without a digital preservation policy, it is expected that the predicted increases in volume of digital information, and the range of formats needing to be preserved will provide the impetus to focus on digital preservation and take practical steps to address its challenges.

**Table 12.** List of open source digital preservation software.

Software detail	GSDL <sup>1</sup>	E-Prints-II <sup>2</sup>	DSpace <sup>3</sup>	Fedora <sup>4</sup>	Koha <sup>5</sup>
Developer	New Zealand Digital Library Project at the University of Waikato	School of Electronics and Computer Science, University of Southampton	Massachusetts Institute of Technology (MIT) and Hewlett-Packard Labs	Cornell University	Katipo Communications
Web Site	<a href="http://www.greenstone.org/">http://www.greenstone.org/</a>	<a href="http://www.eprints.org/">http://www.eprints.org/</a>	<a href="http://www.dspace.org/">http://www.dspace.org/</a>	<a href="http://fedora-commons.org/">http://fedora-commons.org/</a>	<a href="http://koha-community.org/">http://koha-community.org/</a>
Operating System	Unices, Windows	Unices	Unices	Unices, Windows	Linux Windows
Web Server	Apache/ IIS	Apache 1.3	Apache 1.3/2.0 Tomcat	Tomcat 1.4	Apache (2.0 is preferred)
Language	Perl	Mod-Perl 1.0	Java 1.3, JSP	J2SDK v.1.4	Perl
Database	Its own	MySQL	PostgreSQL 7.3	McKoi v.0.94 (uses by default) MySQL/Oracle 9i (optional)	MySQL
Resource	No	OAI Identifiers	CNRI Handles	OAI-PMH	OAI-Server
Identifier			(Similar to URNs)	identifiers (PID)	OAI
Metadata	Dublin Core	Dublin Core	Dublin Core	Dublin Core	Dublin Core

<sup>1</sup>Greenstone Digital Library Software. Retrieved 11.07.2012, from [www.greenstone.org/](http://www.greenstone.org/) Web Site: [http://www.greenstone.org/software details](http://www.greenstone.org/software%20details)

<sup>2</sup>Eprints Digital Library Software. Retrieved 11.07.2012, from <http://www.eprints.org/> Web Site: [http://www.eprints.org/software details](http://www.eprints.org/software%20details)

<sup>3</sup>DSpace Digital Library Software. Retrieved 11.07.2012, from <http://www.dspace.org/> Web Site: [http://www.dspace.org/software details](http://www.dspace.org/software%20details)

<sup>4</sup>Fedora Digital Library Software. Retrieved 11.07.2012, from <http://fedora-commons.org/> Web Site: [http://fedora-commons.org/software details](http://fedora-commons.org/software%20details)

<sup>5</sup>Koha Digital Library Software. Retrieved 11.07.2012, from <http://koha-community.org/> Web Site: [http://koha-community.org/software details](http://koha-community.org/software%20details)

## REFERENCES

- DSpace Digital Library Software. Retrieved 11.07.2012, from <http://www.dspace.org/> Web Site: [http://www.dspace.org/software details](http://www.dspace.org/software%20details).
- Eprints Digital Library Software. Retrieved 11.07.2012, from <http://www.eprints.org/> Web Site: [http://www.eprints.org/software details](http://www.eprints.org/software%20details).
- Fedora Digital Library Software. Retrieved 11.07.2012, from <http://fedora-commons.org/> Web Site: [http://fedora-commons.org/software details](http://fedora-commons.org/software%20details).
- Greenstone Digital Library Software. Retrieved 11.07.2012, from [www.greenstone.org/](http://www.greenstone.org/) Web Site: [http://www.greenstone.org/software details](http://www.greenstone.org/software%20details).
- Gurpreet Singh S (2012). Digitization preserves Punjab region's heritage and culture. Res. Inform. pp.16-17.
- Koha Digital Library Software. Retrieved 11.07.2012, from <http://koha-community.org/> Web Site: [http://koha-community.org/software details](http://koha-community.org/software%20details).
- Market and technology trends analysis (2007). Retrieved 11.07.2012, from [www.digitalpreservationeurope.eu](http://www.digitalpreservationeurope.eu) Web Site: <http://www.digitalpreservationeurope.eu/Publications/dpe-market-analysis.pdf>.
- Meaning of Digital Preservation. Retrieved 18.04. 2012, from [www.dpconline.org](http://www.dpconline.org) Web Site: <http://www.dpconline.org/advice/preservationhandbook/introduction/definitions-and-concepts>.
- Palaiologk Anna (2012). An activity-based costing model for long-term preservation and dissemination of digital research data: the case of DANS International Journal on Digital Libraries. Int. J. Digital Lib 12(4):195-214.
- Robert Neumayer, Open Source in Digital Preservation. Retrieved 24.10.2011, from [www.digitalpreservationeurope.eu](http://www.digitalpreservationeurope.eu) Web Site: [http://www.digitalpreservationeurope.eu/publications/briefs/open\\_source.pdf](http://www.digitalpreservationeurope.eu/publications/briefs/open_source.pdf).
- State University details. Retrieved 06.05.2012, from [www.ugc.ac.in](http://www.ugc.ac.in) Web Site: [http://www.ugc.ac.in/inside/State\\_UniversityNovember2011.pdf](http://www.ugc.ac.in/inside/State_UniversityNovember2011.pdf).