



# National Digital Preservation Programme

Department of Information Technology, Government of India

सत्यमेव जयते

Policy > **Standards** > **Infrastructure** > Sustenance ...



## CENTRE OF EXCELLENCE FOR DIGITAL PRESERVATION

Sponsored by Department of Information Technology

*Preserving Our Digital Universe For Access*



Pune

**Dr. Dinesh Katre**  
Chief Investigator  
Centre of Excellence for Digital Preservation

Associate Director & HOD  
Human-Centred Design & Computing Group  
C-DAC, Pune, INDIA.



United Nations  
Educational, Scientific and  
Cultural Organization



Memory of the World  
20<sup>th</sup> Anniversary



# 4.8 cr pages from govt files untraced

Sharad Vyas & Sandeep Ashar | TNN

Ad day after the Mantralaya inferno, the state government was struggling to track nearly 4.82 crore data pages stored in the 57-year-old building. The records were supposed to be digitised as part of a programme launched in 2010 but could not be because of "the slowness of the process".

Till March this year, officials said, 3.18 crore data pages — or 40% of the total 8 crore leaves — had been scanned and backed up in electronic form. Forming 10 terabyte of data, the pages belonged to records of critical departments, including urban development (UD), housing, home, general administration and environment. Specifically, 12.5 lakh pages from the UD department, 34 lakh from the home department and 37 lakh from the general administration department had been digitised. The process progressed slowly because there was not enough space in the secretariat.

"In the absence of a sepa-



Sharad Vyas

## SAVE AS

According to the state, not all was lost in the Thursday blaze. Owing to a digitisation project that began two years ago, many documents belonging to vital departments were saved in electronic format

Pages scanned until March **3.18 cr**

Total pages to be digitised  
**8 cr**  
(approx)

Files scanned **2.27 lakh**

Cost incurred so far **₹4.5 cr**

Work completed **40%**

Files stored at | Centres in Navi Mumbai and Sion

Files relating to | Urban Development, Home, General Administration, Environment, and Housing departments

Pages yet to be scanned **4.82 cr**

Work still unfinished **60%**

Files relating to | Energy, Transport, Public Works, Industries, Law & Judiciary, and Public Health departments

rate place for scanning documents, we had to carry out the exercise in offices of the departments, which delayed the process," said an official.

Had the digitisation progressed at the right pace,

many of the remaining 4.82 crore pages could probably have been saved. The documents were stored in four record rooms and various departments, which were not adequately fireproofed. "In

hindsight, we could say that the process' pace could have been faster and storerooms better equipped to deal with fire," said another senior official. He added that it would be difficult to predict the state the documents

may be in.

Conspiracy theories of sabotage abounded soon after a blaze broke out at Mantralaya on Thursday afternoon. Many wondered if the fire was started to destroy crucial documents

relating to Adarsh and other high-profile cases. Officials, however, argued that no important data was lost since the major departments were covered in the digitisation programme. Only smaller departments such as transport and tourism were left.

The digitisation was started by the Directorate of Information Technology (DIT) in 2010. A second tender was issued this year and the contract awarded to Navi Mumbai-based Sify Technologies on May 8. For the process, a data centre was created in the new administrative building opposite Mantralaya. But after the inferno, DIT mirrored all 10 terabytes of data on Thursday night and stored in two new centres at Navi Mumbai and Sion.

Chief minister Prithviraj Chavan said on Friday that the state will undertake a massive exercise to rescue files and documents not burnt in the fire and to recreate or restore those that got charred. Nearly 30% (about 1.55 lakh sq ft) of the total built-up area of Mantralaya was affected by the inferno.

(With inputs from Savio D'souza)



## Massive Flooding Damages Several NYC Data Centers

By: Rich Miller  
October 30th, 2012

[Like](#) 292 [Tweet](#) [+1](#) [Share](#) [Print](#)



Flooding from Hurricane Sandy has hobbled two data center buildings in Lower Manhattan, taking out diesel fuel pumps used to refuel generators. A third building at 121 Varick is also reported to be without power. There were also reports of outages for some tenants at a major data hub at 111 8th Avenue, and many other New York area facilities were running on generator power amid widespread utility outages.



Connect via Twitter, Facebook, LinkedIn & RSS.

Server Solutions  
for the Precision-Fit  
Data Center

[LEARN MORE >](#)

ZT 1200RI Servers featuring the  
NEW Intel® Xeon® processor E5-2600



### Featured Cloud Articles

**Dell Partners With Morphlabs on 'Private Cloud in a Box'** Dell is partnering with Morphlabs to bring an integrated cloud appliance to market as part of Dell's new Emerging Solutions Platform Partner Program.



**Understanding Cloud APIs Now,** administrators can integrate applications and other workloads into cloud APIs. The question is which model is right and how can you incorporate an API into your cloud?



"The Mission Critical Design Assist Program™"

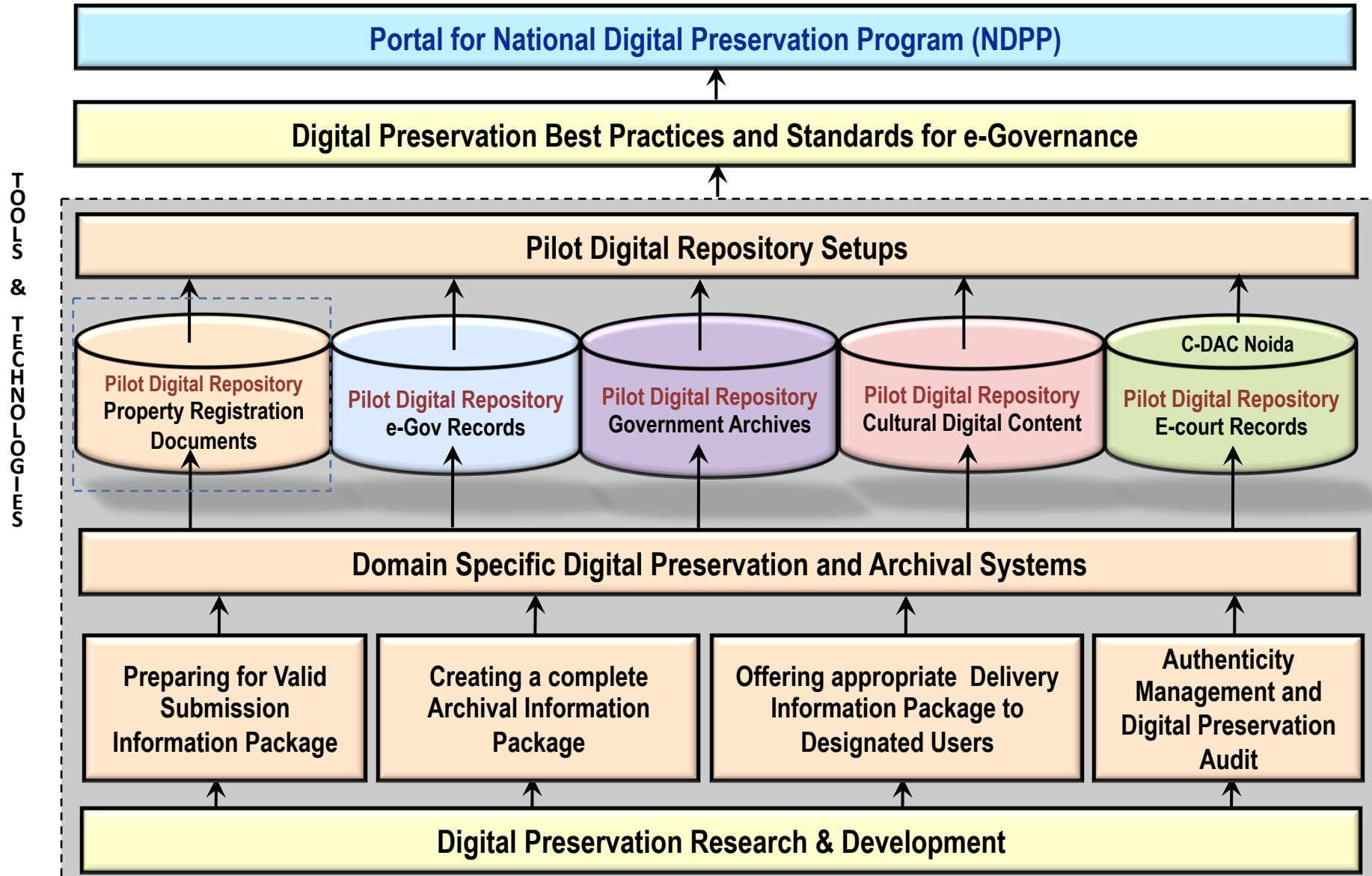


- Indirect - Direct Cooling Solutions
- Free Cooling Strategies
- Significantly Reduced PUE

www.unitedmetal.com

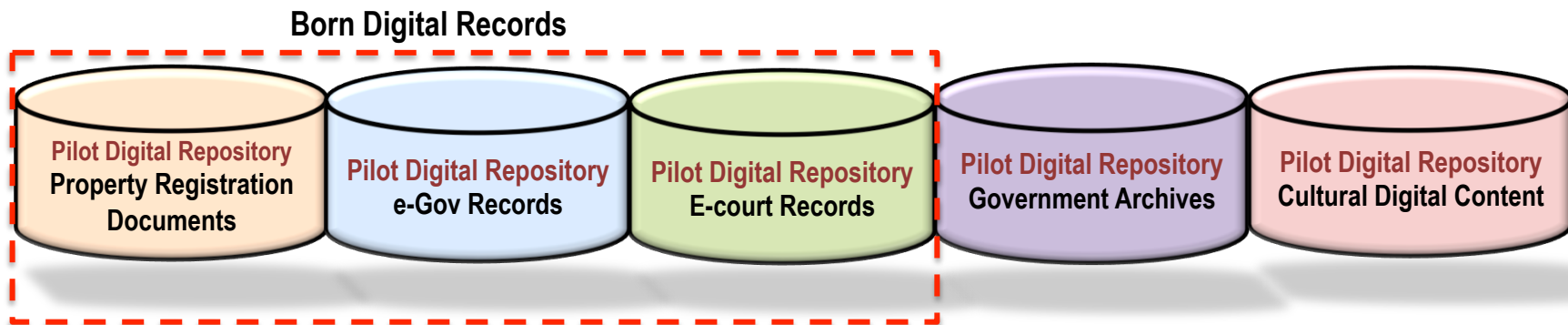


### Scope of CoE-DP Project





**Clarity on what needs to be preserved**  
**Legacy data not in condition for preservation**  
**Process definition for new data**



### Pre-SIP Processing

#### Data alignment activity

- format conversion
- schema design
- Mapping of datatable with schemas

#### Creation of SIPs

- Digital object formation
- XML generation
- Resolve PKI issues

#### Batch processing of SIPs

#### Process alignment for e-records production



# National Digital Preservation Programme

Department of Information Technology, Government of India

सत्यमेव जयते



**CENTRE OF EXCELLENCE FOR  
DIGITAL PRESERVATION**

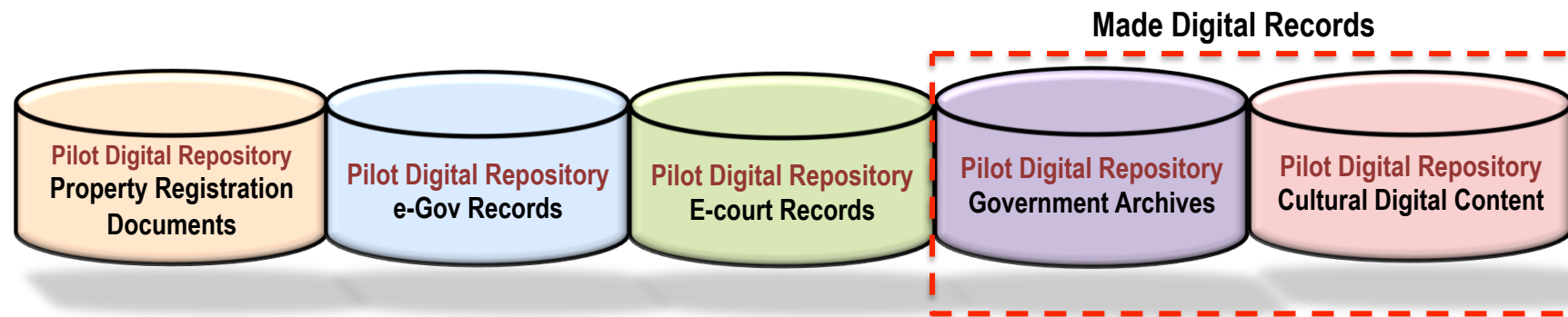
Sponsored by Department of Information Technology

Preserving Our Digital Universe For Access



Pune

**Digital objects and cataloging database are not linked**  
**Closed / obsolete file formats**  
**Non-refreshed CDs / DVDs**  
**Digital authenticity is not addressed**



## Pre-SIP Processing

Data alignment activity

- format conversion
- mapping with metadata standards
- assign object IDs
- XML generation

Complete the SIPs as per the template

Quality assurance for SIPs

Process alignment



National Digital Preservation Programme  
Department of Information Technology, Government of India



CENTRE OF EXCELLENCE FOR  
**DIGITAL PRESERVATION**  
Sponsored by Department of Information Technology  
*Preserving Our Digital Universe For Access*



Nos

Software development initiatives

References, Standards & Best Practices

SD1



NDSAP - 2012  
IT ACT 2000  
SPIDeR, METS, DCC, MODS  
PREMIS  
ISO 15489 Record Management  
ISO/IEC 27000 Information Security Management  
ISO 14721:2003 OAIS

SD2



NARA Digitization Guidelines  
DCC and METS  
ISO 14721:2003 OAIS  
ISO 13028 – Guidelines for digitization of records

SD3



ISO 25577 MarcXchange / MARC21  
Standards related to Audio, Video and 3D data  
DCC, METS, MODS  
ISO 14721:2003 OAIS

SD4

Webportal for National Digital Preservation Programme

[www.ndpp.in](http://www.ndpp.in)

Login as Director / Repository Administrator / Archivist / SIP Creator/ Scholar

LOGIN AS



Director



Repository Administrator



Archivist



SIP Creator



Scholar






# National Digital Preservation Programme

Department of Information Technology, Government of India



**CENTRE OF EXCELLENCE FOR  
DIGITAL PRESERVATION**  
Sponsored by Department of Information Technology  
*Preserving Our Digital Universe For Access*






Centre of Excellence for  
Digital Preservation  
Sponsored by Department of  
Information Technology,  
Government of India

National Digital Preservation Programme

## Digital Preservation Estimation tool Beta

Designed & Developed By



**YOUR SEARCH QUERY:** Welcome **Guest** Logout

SELECT THE CATEGORY FOR RETENTION RULES

Retention rules Substantive Functions Of Ministry Of Labour & Employment Directorate General Of Employment & Training ▼

Section Name	<div style="border: 1px solid #ccc; padding: 2px;"> <p>ALL ▼</p> <p>Select</p> <p style="background-color: #007bff; color: white;">ALL</p> </div>
Duration options	<ul style="list-style-type: none"> <li>Advance Vocational Training Schemes/Technical Stores (AVTS) Section</li> <li>Apprenticeship Programme (AP) Section</li> <li>Central Employment Exchange (CEE) Section</li> <li>Central Project Implementation Unit (CPIU)</li> <li>Curricular Development (CD) Section</li> <li>Coordination Section</li> <li>Employment Exchange-I (EE-I) Section</li> <li>Employment Exchange-II(EE-II) Section</li> <li>Employment Marketing Information (EMI) Division.</li> <li>Evaluation And Implementation (E&amp;I) Section</li> <li>Vigilance &amp; Foreign Training Assignments (VFTA) Section</li> <li>Man Power (General) Section</li> <li>Planning Cell Training (PCT) Section</li> <li>Planning Cell (Employment) (PCE) Section</li> <li>Statistical Section</li> <li>Survey &amp; Study Division</li> <li>Training Craftsmen (TC-I) Section</li> <li>Training Craftsmen - Desk (TC-II) Section</li> </ul>
Keyword	<input style="width: 100%;" type="text"/>

Designed and developed by HCDC, C-DAC

## Digital Preservation Estimation Tool

### Features:

- Select retention rules
- Search record types
- Select and quantify
- Define format specs
- Define AIP specs
- Produce storage estimate
- Produce cost estimate
- Produce time estimate



YOUR SEARCH QUERY:

Welcome **Guest** [Logout](#)

RETENTION PERIOD

All retention years

SEARCH RESULT

Sort by

DETAILED VIEW

NO. OF RECORDS FOR ESTIMATION : 1739

FIRST

<< PREV

174 / 174

NEXT >>

LAST

NO.	RECORDS NAME	RETENTION PERIOD	<input type="checkbox"/>
1731	Adaptation of general Aptitude Test Battery : Modification	25 year	<input checked="" type="checkbox"/>
1732	Preparation of software in respect of Educational and Vocational facilities available for the unemployed youth.	25 year	<input type="checkbox"/>
1733	Training Programme: Vocational/Industrial Training in Rural Administration etc.: Schemes under various Training Programmes	25 year	<input type="checkbox"/>
1734	Plan Schemes - National/ International : Opening of Vocational Training Institutions/Model Training Institutions	25 year	<input checked="" type="checkbox"/>
1735	Advance Vocational Training Schemes : Policy matters regarding the project.i.e. Project Agreements/ Expansion/Introduction of Projects/Assistance of ILO in formulation of the Projects/Development Programmes and their implementations.	25 year	<input type="checkbox"/>
1736	Pay Bill register.	35 year	<input checked="" type="checkbox"/>
1737	Office copies of Establishment pay bills and related schedules (in respect of period for which pay bill register is not maintained).	35 year	<input type="checkbox"/>
1738	Pay Bill Register	35 year	<input type="checkbox"/>
1739	Office copies of establishment pay bills and related schedules (in respect of period for which Pay Bill Register is not maintained).	35 year	<input checked="" type="checkbox"/>

BACK TO SEARCH

VIEW ESTIMATION LIST



YOUR SEARCH QUERY:

Welcome Guest [Logout](#)

RETENTION PERIOD

All retention years

SEARCH RESULT

Sort by

DETAILED VIEW

NO. OF RECORDS FOR ESTIMATION : 1739

FIRST

<< PREV

40 / 174

NEXT >>

LAST

NO.	RECORDS NAME	RETENTION PERIOD	<input type="checkbox"/>
391	General staff welfare measures : Broad aspects	Permanent	<input checked="" type="checkbox"/>
392	General staff welfare measures : CSS (Recognition of Service Association) Rules	Permanent	<input type="checkbox"/>
393	General staff welfare measures : Recognition of Association (individual cases)	Permanent	<input type="checkbox"/>
394	Departmental council / office council : General aspects/ instructions	Permanent	<input checked="" type="checkbox"/>
395	Departmental council / office council : Staff Union/Association : Recognition	Permanent	<input type="checkbox"/>
396	Grants-in-aid : General aspects	Permanent	<input checked="" type="checkbox"/>
397	Creation and classification of posts : Revision of scales of pay	Permanent	<input type="checkbox"/>
398	Co-operative Societies : Rules and bye-laws (general aspects)	Permanent	<input type="checkbox"/>
399	Recruitment : Recruitment (general aspects) including provisions of the Constitution	Permanent	<input type="checkbox"/>
400	Recruitment : UPSC (Exemption from Consultation) Regulations	Permanent	<input type="checkbox"/>

[BACK TO SEARCH](#)

[VIEW ESTIMATION LIST](#)



ESTIMATION LIST:

Welcome **Guest** [Logout](#)

RETENTION PERIOD WISE STORAGE : ALL		<input type="checkbox"/> DETAILED VIEW	NO. OF RECORDS FOR ESTIMATION : 7		FIRST	« PREV	1 / 1	NEXT »	LAST
<input type="checkbox"/>	RECORDS NAME	SETTINGS	NO. OF FILES	PAGES (AVG/DOC)	SIZE(GB)	RETENTION PERIOD	REMOVE		
<input type="checkbox"/>	General staff welfare measures : Broad aspects		<input type="text" value="53"/>	<input type="text" value="132"/>	100.78 <a href="#">CALCULATE</a>	Permanent			
<input type="checkbox"/>	Grants-in-aid : General aspects		<input type="text" value="40"/>	<input type="text" value="145"/>	83.55 <a href="#">CALCULATE</a>	Permanent			
<input type="checkbox"/>	Office copies of establishment pay bills and related schedules (in respect of period for which Pay Bill Register is not maintained.		<input type="text" value="145"/>	<input type="text" value="167"/>	87.21 <a href="#">CALCULATE</a>	35 year			
<input type="checkbox"/>	Adaptation of general Aptitude Test Battery : Modification		<input type="text" value="578"/>	<input type="text" value="154"/>	320.57 <a href="#">CALCULATE</a>	25 year			
<input type="checkbox"/>	Plan Schemes - National/ International : Opening of Vocational Training Institutions/Model Training Institutions		<input type="text" value="445"/>	<input type="text" value="200"/>	320.53 <a href="#">CALCULATE</a>	25 year			
<input type="checkbox"/>	Departmental council / office council : General aspects/ instructions		<input type="text" value="45"/>	<input type="text" value="47"/>	30.47 <a href="#">CALCULATE</a>	Permanent			
<input type="checkbox"/>	Pay Bill register.		<input type="text" value="167"/>	<input type="text" value="157"/>	94.43 <a href="#">CALCULATE</a>	35 year			

[ADD MORE RECORDS](#) [ADD CUSTOM RECORDS](#) [NEW LIST](#)



ESTIMATION LIST:

Welcome **Guest** [Logout](#)

RETENTION PERIOD WISE STORAGE :   DETAILED VIEW NO. OF RECORDS FOR ESTIMATION : 7

<input type="checkbox"/>	RECORDS NAME	SETTINGS	NO. OF FILES	PAGES (AVG/DOC)	SIZE(GB)	RETENTION PERIOD	REMOVE
<input type="checkbox"/>	General staff welfare measures : Broad aspects		<input type="text" value="53"/>	<input type="text" value="132"/>	100.78 <input type="button" value="CALCULATE"/>	Permanent	
<input type="checkbox"/>	Grants-in-aid : General aspects		<input type="text" value="40"/>	<input type="text" value="145"/>	83.55 <input type="button" value="CALCULATE"/>	Permanent	
<input type="checkbox"/>	Office copies of establishment pay bills and related schedules (in respect of period for which Pav Bill Register is not		<input type="text" value="145"/>	<input type="text" value="167"/>	87.21 <input type="button" value="CALCULATE"/>	35 year	

STORAGE REQUIRED: TOTAL SIZE OF CALCULATED RECORDS: 1.01 TB SIZE IN :

MEDIA	CAPACITY	NUMBER
CD	700 MB	1518
DVD (SINGLE LAYER)	4.7 GB	221
DVD (DOUBLE LAYER)	8.5 GB	123
BLU-RAY (SINGLE LAYER)	25 GB	42
BLU-RAY (DOUBLE LAYER)	50 GB	21
TAPE (DAT 72)	36 GB	29



Digital Preservation Estimation Tool-Storage Requirement

No.of Records for Archival	7
Size of Records	1.01 TB / 1037.55 GB

S.No	Record Name	Retention Period	Size (GB)
1	General staff welfare measures : Broad aspects	Permanent	100.78
2	Grants-in-aid : General aspects	Permanent	83.55
3	Office copies of establishment pay bills and related schedules (in respect of period for which Pay Bill Register is not maintained.	35 year	87.21
4	Adaptation of general Aptitude Test Battery : Modification	25 year	320.57
5	Plan Schemes - National/ International : Opening of Vocational Training Institutions/Model Training Institutions	25 year	320.53
6	Departmental council / office council : General aspects/ instructions	Permanent	30.47
7	Pay Bill register.	35 year	94.43

Storage media Required

MEDIA	CAPACITY	REQUIRED
CD	700 MB	1518
DVD (SINGLE LAYER)	4.7 GB	221
DVD (DOUBLE LAYER)	8.5 GB	123
BLU-RAY (SINGLE LAYER)	25 GB	42
BLU-RAY (DOUBLE LAYER)	50 GB	21



# National Digital Preservation Programme

Department of Information Technology, Government of India



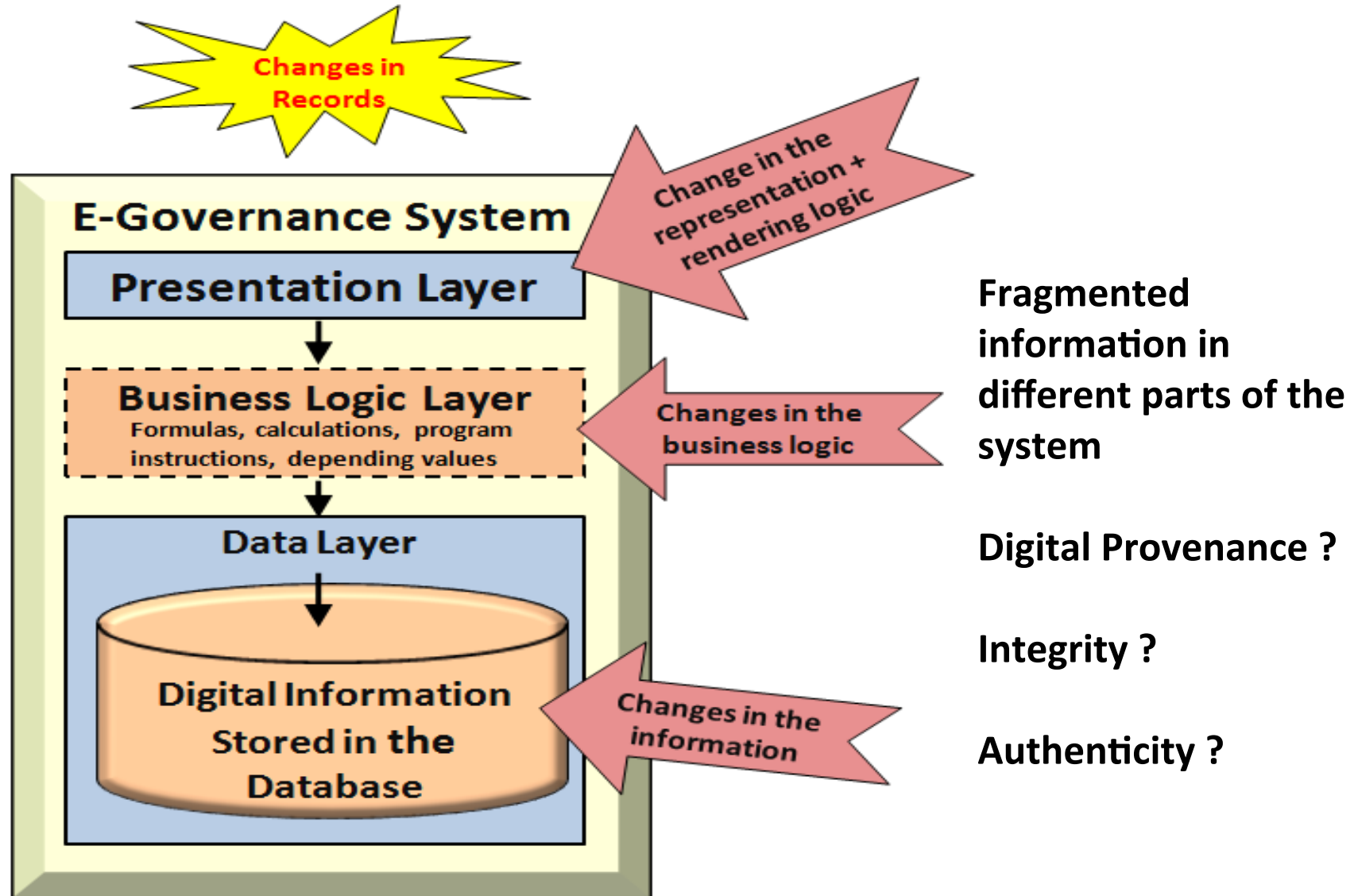
**CENTRE OF EXCELLENCE FOR  
DIGITAL PRESERVATION**

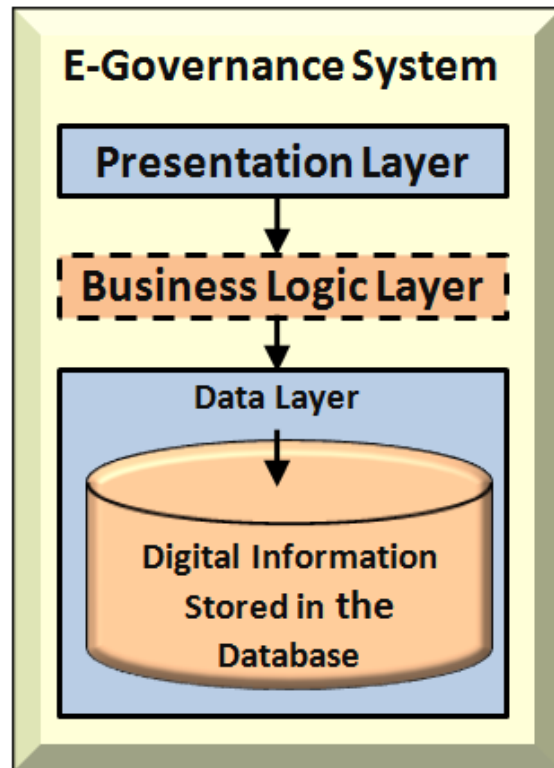
Sponsored by Department of Information Technology

Preserving Our Digital Universe For Access



Pune





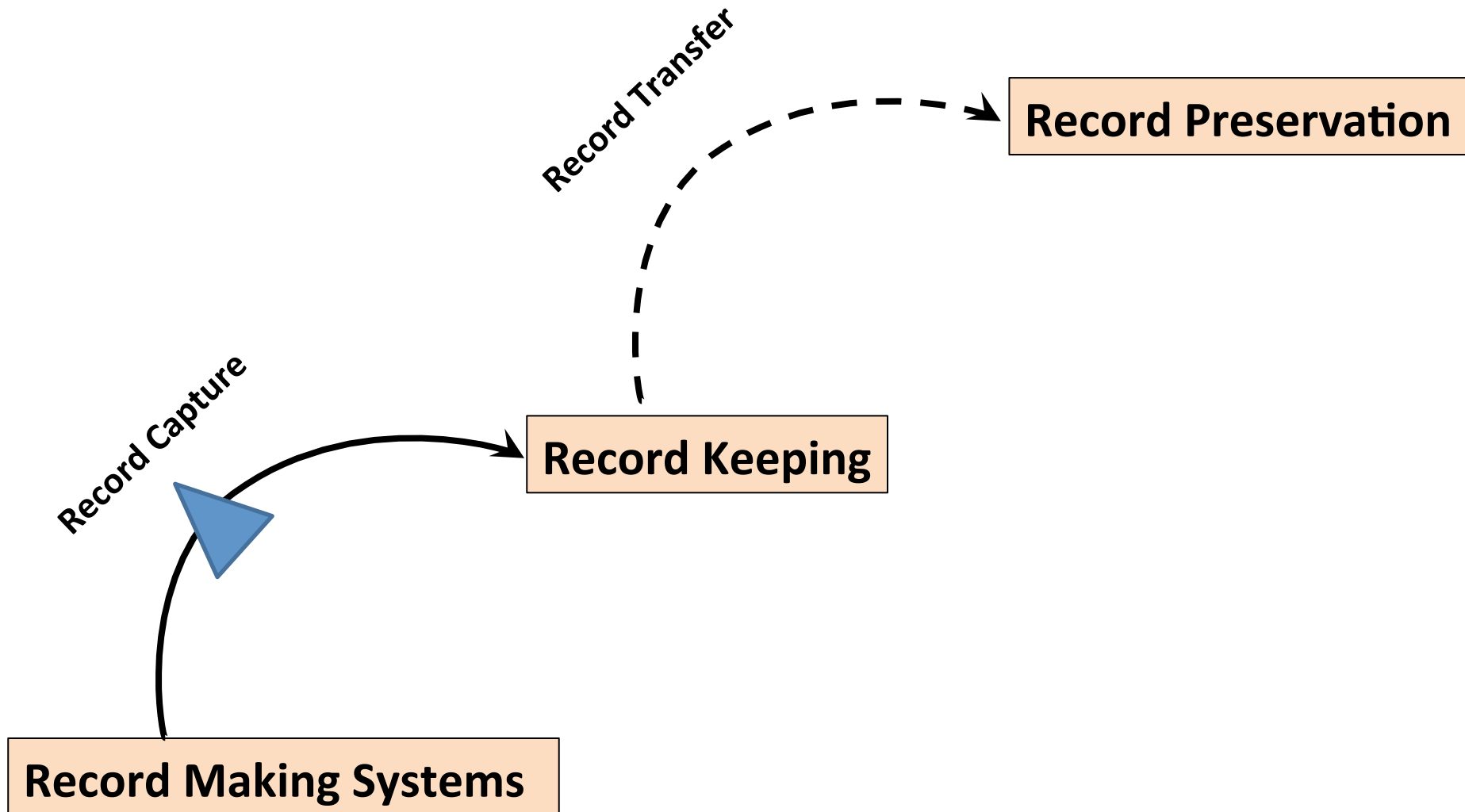
## Types of Database Records

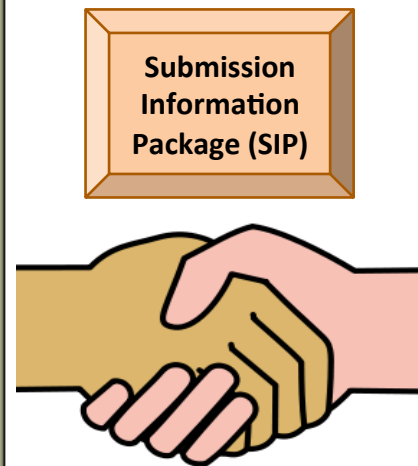
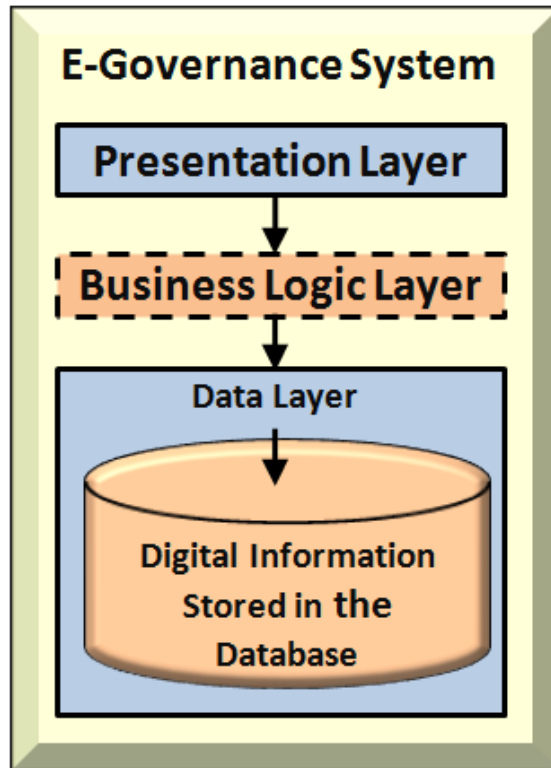
- Fixed information / values / content stored in the database – represented as a document
- Calculated values derived from the database using the business logic – represented as a document
- Incrementally growing transactional records (transactions reach the end of life cycle but the record continues to evolve)



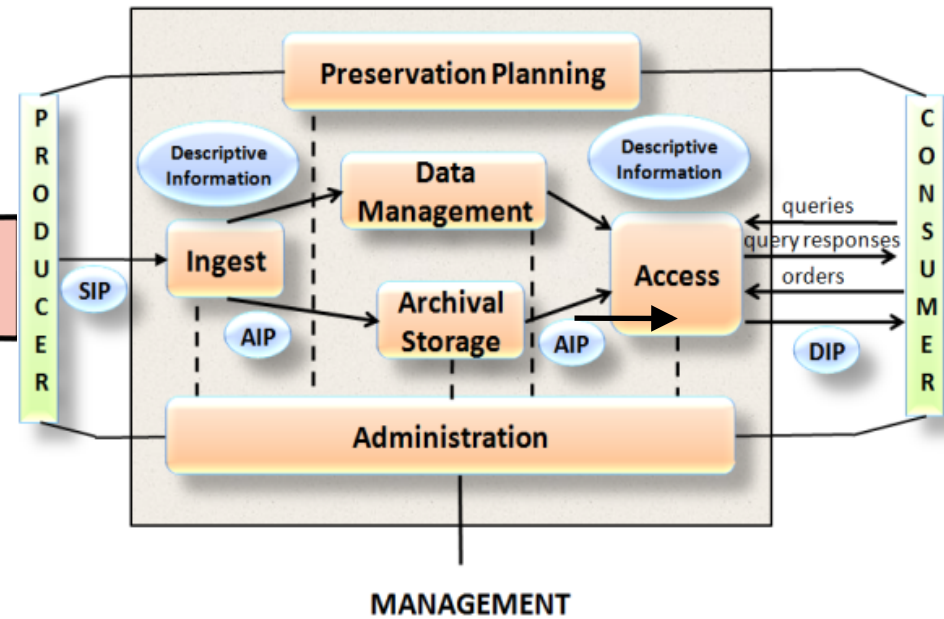


## E-record Management Lifecycle: Indian e-Governance Scenario





### Open Information Archival System ISO 14721:2003





National Digital Preservation Programme  
Department of Information Technology, Government of India



CENTRE OF EXCELLENCE FOR  
**DIGITAL PRESERVATION**  
Sponsored by Department of Information Technology  
*Preserving Our Digital Universe For Access*



**Draft standards developed during 2011-2012**

# **Production of Preservable e-Records (PROPeR: Best Practices & Guidelines)**

Version 0.1

- Fixed digital object
- Consistent and logical name
- Archival format specification
- Logical and conceptual representation
- Authenticity
- Preservation information



National Digital Preservation Programme  
Department of Information Technology, Government of India



CENTRE OF EXCELLENCE FOR  
**DIGITAL PRESERVATION**  
Sponsored by Department of Information Technology  
*Preserving Our Digital Universe For Access*



**Draft standards developed during 2011-2012**

## **e-Governance Standard for Preservation Information Documentation of Electronic Records**

Version 0.1

**(e-Gov SPIDeR: Metadata Dictionary & Schema)**

- Cataloging information
- Enclosure information
- Provenance information
- Fixity information
- Representation information
- Digital signature information
- Access rights information



# Roadmap

2015

**DIGITALAYAYA**™

Replicate  
Trustworthy Digital Repositories

Develop the  
competencies  
for Audit and  
Certification -  
ISO 16363

Establish Pilot  
Digital  
Repositories

Conduct  
Training  
Certify the  
DROs,  
Archivists

Curriculum  
for Digital  
Preservation  
Training

Define Digital  
Preservation  
Standards  
and Best  
Practices

Develop the  
Digital  
Archival &  
Preservation  
System(s)

Draft the  
National  
Digital  
Preservation  
Policy

Recommendations for  
National Digital Preservation Programme  
(Foundation provided by the study report)



# CENTRE OF EXCELLENCE FOR DIGITAL PRESERVATION

Sponsored by Department of Information Technology  
*Preserving Our Digital Universe For Access*



Pune

A+ A A-



search...

## Main Menu

- National Study Report
- Centre of Excellence
  - Vision
- TechWatch
  - Long Term File System(LTFS)
- Tools & Technologies
- Digital Preservation Standards
- Partner Organizations
- Events
  - ICLAM, 2011
  - National Meet, 2010
  - Indo-Us Workshop, 2009
  - National Workshop, 2008
- Research Publications
- Contact Us

## Resources

- Photo Gallery

Home ▶ Partner Organizations

Tuesday, 05 April 2011 12:13

### Our Partner Organizations



[Centre for Development of Advanced Computing \(C-DAC\), Pune](#)

**Centre of Excellence for Digital Preservation**



[National e-Governance Action Plan \(NeGP\), Department of Information Technology, New Delhi](#)



[National Informatics Centre, Lucknow](#)



[National Archives of India, New Delhi](#)



[Indira Gandhi National Centre for Arts, New Delhi](#)



[Centre for Development of Advanced Computing \(C-DAC\), Noida](#)



[Delhi High Court, New Delhi](#)

## Login

Username

Password

Remember Me

➤ [Forgot password?](#)

➤ [Forgot username?](#)

## Polls

**Do the government/ private organizations have a documented preservation policy for the archival of the records generated by them?**

- Yes, we follow the preservation policy.
- It is available but not practiced.



National Digital Preservation Programme  
Department of Information Technology, Government of India



CENTRE OF EXCELLENCE FOR  
**DIGITAL PRESERVATION**  
Sponsored by Department of Information Technology  
*Preserving Our Digital Universe For Access*

सी डैक  
CDAC  
Pune

Thank You 😊



**National Digital Preservation Programme**  
**[www.ndpp.in](http://www.ndpp.in)**