Digital Preservation

measuring our capability and 'confronting the abyss'

Sarah Slade

on behalf of the NSLA Digital Preservation Group

NSLA Digital Preservation Group

- Gaining a shared understanding of current digital collection management practices and workflows in NSLA libraries

- Sharing information about digital preservation best practice

- Determining the core requirements for managing the preservation of digital collections in NSLA libraries and identifying opportunities for collaboration

RE
IMAGINING
LIBRARIES

Work Packages

Work Package 1: What is it and why?

Work Package 2: How well?

Work Package 3: Who?

Work Package 4: Nuts & Bolts

Work Package 5: Collaboration & Partnership

Work Package 6: Confronting the Abyss

RE
IMAGINING
LIBRARIES

Work Packages

Work Package 1: What is it and why?

Work Package 2: How well?

Work Package 3: Who?

Work Package 4: Nuts & Bolts

Work Package 5: Collaboration & Partnership

Work Package 6: Confronting the Abyss

RE
IMAGINING
LIBRARIES

Work Packages

Work Package 1: What is it and why?

Work Package 2: How well?

Work Package 3: Who?

Work Package 4: Nuts & Bolts

Work Package 5: Collaboration & Partnership

Work Package 6: Confronting the Abyss

RE
IMAGINING
LIBRARIES

How well?

Who?

Digital Preservation Environment Maturity Matrix

Digital Preservation Capability Maturity Matrix

A matrix outlining the twenty components of an ideal digital preservation system, and marking stages of development and maturity for each NSLA library.

A matrix to demonstrate organisational capability in each NSLA library

RE IMAGINING LIBRARIES

Aim was to …

- Identify where libraries currently sit against the list of components

- Help the libraries to identify development needs

- Help NSLA identify collaboration and development needs

RE
IMAGINING
LIBRARIES

Assumptions

The Digital Preservation project assumes that each Library?
Is actively collecting digital materials, both born and digitised
Is committed to preserving its digital materials for the long term
Has resources (including staff or vendor with appropriate skills) dedicated to the task
Has a sustainable funding model
Aims to comply with the OAIS responsibilities

RE
IMAGINING
LIBRARIES

Open Archival Information System

The reference model for an open archival information system (OAIS) is a commonly accepted standard amongst the digital preservation community

The OAIS standard is a high-level abstract model, which …

Provides a framework, including terminology and concepts, for describing and comparing architectures and operations of existing and future archives

1. Pre-ingest Activities

2. Ingest

3. Archival Storage

4. Data Management

5. Administration

6. Digital Preservation Planning

7. Access

How well?

1. Pre-ingest Activities

What system policies and standards related to digital collecting and preservation do you have in place in your library?

2. Ingest

What SIPs do you receive from producers, and how?

How do you validate the SIPs?

How do you generate AIPs from SIPs?

What metadata do you extract from AIPs or collect from other sources, and how?

3. Archival Storage

How are your AIPs stored?

What proactive measures do you take to refresh your archival media/storage?

What routine and special error checking do you perform to make sure that no components of the AIP are corrupted in archival storage or during any internal archival storage data transfers?

What IT disaster recovery plans and business continuity plans does your library have in place to protect your digital assets?

4. Data Management

How do you store, maintain and update metadata for your library's digital collection content?

How do you monitor collection status?

5. Administration

How do you negotiate submission agreements and audit submissions to ensure that they meet your institution's standards?

How do you manage system configuration?

What mechanisms do you provide to restrict or allow physical access to elements of the archive, as determined by archive policies?

How do you establish and maintain system standards and policies?

6. Digital Preservation Planning

How do you monitor changes in the Digital Preservation and ICT technology environments and in the designated community's service requirements and knowledge base?

How do you develop preservation strategies and standards?

How do you develop packaging designs and preservation actions plans?

7. Access

How do you provide access to your data?

How do you ensure that the user is authorised to access and receive the requested items?

Who?

1. Pre-ingest Activities

What capability does your Library have for pre-ingest activities?

2. Ingest

What capability does your Library have for Ingest processes?

What capability does your Library have for validating SIPs?

Does your Library routinely generate AIPs from SIPs?

What capability does your Library have for metadata management and related processes?

3. Archival Storage

What capability does your Library have for the management of storage systems and related processes?

What capability does your Library have for the management of refresh processes associated with storage systems?

What capability does your Library have to ensure data corruption does not occur on the storage system or during data transfers?

What capability does your Library have to develop IT disaster recovery plans and business continuity plans and to implement them if required?

4. Data Management

What capability does your Library have to store, maintain and update metadata for your digital collection content?

What capability does your Library have to monitor collection status?

5. Administration

What capability does your Library have for the negotiation of submission agreements and to audit submissions to ensure that they meet your institution's standards?

What capability does your Library have to manage system configuration processes?

What capability does your Library have to manage physical system security processes?

What capability does your Library have to establish and maintain system standards and policies?

6. Digital Preservation Planning

What capability does your Library have to monitor changes in the Digital Preservation and ICT technology environments and in the designated community's service requirements and knowledge base?

What capability does your Library have to develop preservation strategies and standards?

What capability does your Library have to develop packaging designs and preservation actions plans?

7. Access

What capability does your Library have in providing access to your data?

What capability does your Library have to ensure the user is authorised to access and receive the requested items?

IMAGINING
LIBRARIES

3. Archival Storage

How are your AIPs stored?
[*This may include removing objects from media, on-site control vs. contractual off-site control, choosing appropriate storage, etc.*]

What proactive measures do you take to refresh your archival media/storage?
[*This may include copying the files onto new media (which could be the same type of media or another technology) or the replacement of a storage system (including SAN, NAS or other spinning disk or solid state storage system)*
*Note: archival media can be CDs, tapes, spinning discs, etc.*]

What routine and special error checking do you perform to make sure that no components of the AIP are corrupted in archival storage or during any internal archival storage data transfers?
[*This may include integrity checking, cyclic redundancy checking, etc.*]

What IT disaster recovery plans and business continuity plans does your Library have in place to protect your digital assets?
[*This may include multiple copies, backups, geo-separate locations, un-interruptible power supply, etc.*]
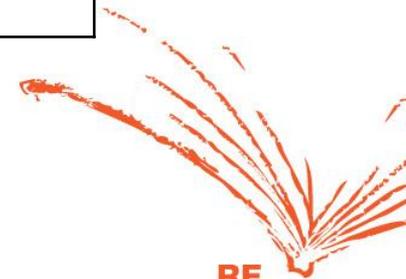
1. Pre-ingest Activities

What capability does your Library have for pre-ingest activities?

To assist in determining your response:
Consider these scenarios:
Scenario 1: An opportunity arises to collect the pre-press files from a regional newspaper.
How would your library handle this, is it obvious which teams/staff needs to be involved, is it a well-managed smooth process with all the relevant teams involved or a 'special' situation requiring special meetings and one-off or new processes to be established? Is the task handed to IT with little oversight from the business?

Scenario 2: Your library has a Flickr group and appraises contributions for addition to the Library's collections.
Is this a straight-forward 'business as usual' process or does it depend on one person remembering to do it resulting in an ad hoc approach, with most other staff knowing nothing about it?

Scenario 3: Donations of material contain a mix of paper and media such as diskettes, CD-ROMs, USBs etc. Do staff readily appraise the material on media or do they often seek support, such as approaching IT, for help on a case by case basis? Do they prefer that 'some-one else' handles the digital material?

Consider these questions:
- How many positions does your Library have dedicated to these functions or is this work tacked onto existing roles?
- What happens when key individuals are not available?
- Is the work articulated into Position or Role Descriptions?
- Does the Library recognise that the actions taken in response to the scenarios above, are Digital Collecting and Digital Preservation Pre-Ingest activities?
- Are the activities talked about using the language of digital preservation?
- Is there concern expressed about the level of resourcing/capability in this area at your Library?
- Is improved digital preservation capability highlighted as a requirement through workforce planning processes?
- Do the Library's HR and management processes help/hinder greater capability in digital preservation?

Note:
Pre-ingest activities include
- *Defining the designated community*
- *Consulting and communicating with the designated community*
- *Creating, maintaining and publishing policies and guidelines relating to digital collecting*
- *Determining degree of compliance with standards and defining local standards*
- *Negotiating with Producers for preferred content formats and rights*
- *Documenting and maintaining agreements*
- *Assisting Producers to submit content and undertaking troubleshooting when necessary*
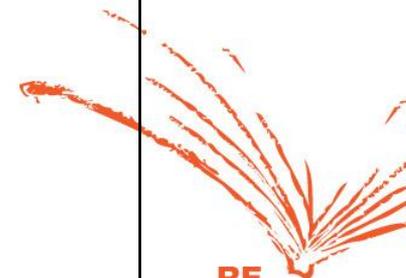
RE
IMAGINING
LIBRARIES

Maturity Models

Capability Maturity Matrix
Level 1 - Initial
Level 2 - Repeatable
Level 3 - Defined
Level 4 - Managed
Level 5 - Optimising

People - Capability Maturity Matrix
Level 1 - Initial
Level 2 - Managed
Level 3 - Defined
Level 4 - Predictable
Level 5 - Optimising

Level 1 - Initial
Processes are usually ad hoc.

Level 2 – Repeatable
Digital preservation achievements are repeatable, but the processes may not repeat for all digital preservation activities in the organisation.

Level 3 – Defined
Digital preservation activities are performed and managed according to documented plans.

| Level 4 – Managed |
|---|
| Management can effectively control the digital preservation effort, using precise measurements. |

Level 5 – Optimising
At this level the organisation focuses on continually improving process performance through both incremental and innovative technological improvements.

Description	Comments about current state and/or future plans	Current CMM Self Rating[1]	Planned CMM Self Rating [2]
1. Pre-ingest Activities			
What system policies and standards related to digital collecting and preservation do you have in place in your Library? [*This may include description of designated communities, collection development policy, digital preservation policy, evaluation of content, submission guidelines and agreements (including rights management)*]			
2. Ingest			
What SIPs do you receive from producers, and how? [*This may include solicited, unsolicited and ad-hoc submissions from external as well as internal producers via ftp site, CD/DVD carriers, email, systematic workflow, etc.*]			
How do you validate the SIPs? [*This may include assessing and accepting content and formats, validating composition of SIP, fixity checks, checking for TPMs, validating metadata, etc.*]			
How do you generate AIPs from SIPs? [*This may include re-arrangement and description of content, creating unique & persistent identifier, structural map, etc.*]			
What metadata do you extract from AIPs or collect from other sources, and how? [*This may include:* • *technical metadata (such as format identification and validation)* • *descriptive metadata* • *administrative metadata*]			

[1] Please insert your CMM self rating (assessed level from 1 to 5) for current state digital preservation maturity to 30 June 2013

[2] Please insert your CMM self rating (assessed level from 1 to 5) for planned digital preservation maturity (to 30 June 2016)
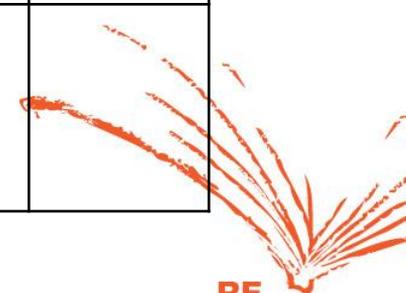
RE IMAGINING LIBRARIES

Aim was to …

- Identify where libraries currently sit against the list of components

- Help the libraries to identify development needs

- Help NSLA identify collaboration development and needs

RE
IMAGING
LIBRARIES

Demonstrated a clear picture of the state of digital preservation in NSLA libraries that were valid and useful

Doesn't allow meaningful comparison between individual libraries

Can be used to demonstrate development over time

Does allow people to understand what is needed and what the steps are in an accessible way

RE
IMAGINING
LIBRARIES

Next steps

- Combine the two matrices into a single Digital Preservation Maturity Matrix

- Develop a guidance sheet for allocation of the maturity ratings to improve comparisons over time

- Repeat the combined tool every three years and track developments and areas in need of investment

- Improve skills and capability through the work of the NSLA Digital Skills Group that was created in response to the findings from the work of the Digital Collecting Group and the initial findings from these Digital Preservation survey tools

- Broaden the use of the survey tool into other GLAM sectors …

RE
IMAGINING
LIBRARIES

Work Packages

Work Package 1: What is it and why?

Work Package 2: How well?

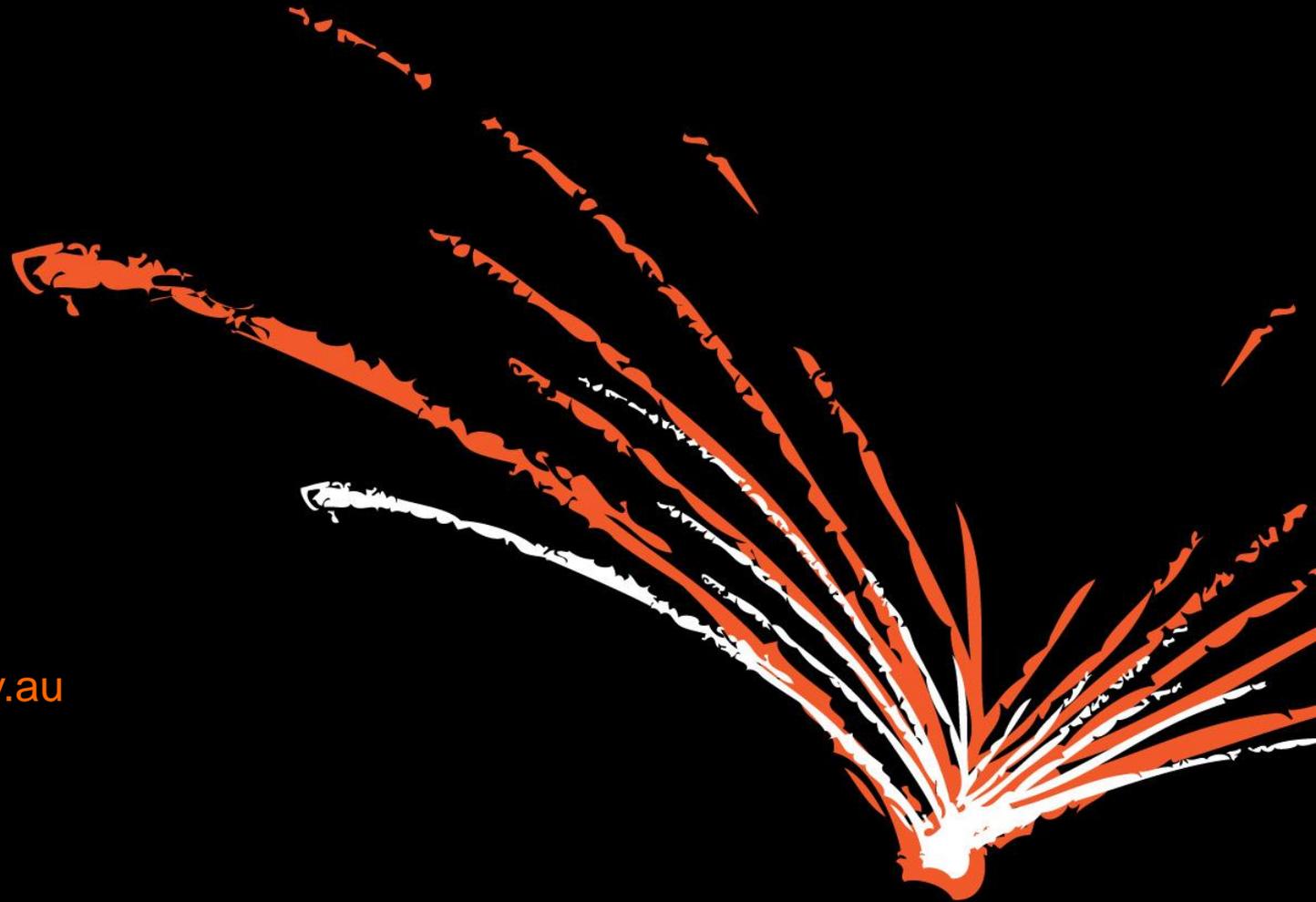Work Package 3: Who?

Work Package 4: Nuts & Bolts

Work Package 5: Collaboration & Partnership

Work Package 6: Confronting the Abyss

RE
IMAGINING
LIBRARIES

Thank you

sslade@slv.vic.gov.au