

# Migration Methodology

The Digital Archives Migration Methodology supports the transfer of digital records from NSW Government agencies to the Digital State Archive. Rather than adopting a single approach for all such transfers, State Records NSW defines custom migration plans to suit the particular requirements of different sets of records. Each transfer is managed as a separate project. The methodology is a framework to guide these projects.

By blending project management and data migration techniques, the Digital Archives Migration Methodology provides a structured and planned approach to each migration project. It also permits flexibility for dealing with many types of migrations, from very simple ones to complex ones involving many record types and stakeholders.

## 1. Project Planning Phase



The Project Planning phase establishes a framework for the migration project. The purpose of this phase is to define the project goals and identify stakeholders, risks, and resources. The depth of planning required in this phase will vary depending on the complexity of the project. For example, a project involving the transfer of a single audio file might be very small in comparison with a project involving the transfer of a business or email system. The key deliverable of this phase is the project plan.

## 2. Migration Planning Phase

The goal of the Migration Planning phase is to develop the migration plan. The migration plan is a document that identifies and documents the activities to be carried out during the migration of a recordkeeping system into the Digital State Archive. These activities include file format migration, metadata mapping, and data transformation. Like project plans, migration plans are tailored to suit the requirements of the particular project. But this doesn't mean re-inventing the wheel each time; the decisions and lessons learned in each project are documented and can be re-used in subsequent projects.

## 3. Migration Phase

During Migration phase, the migration plan is executed. It is in this stage that any necessary preservation activities are performed.

## 4. Project Closure Phase

The Project Closure phase closes the project and identifies any required post-project activities (such as the disposal of source records).

### Tools



**Basecamp** *Adopted*  
<https://basecamp.com>  
 Basecamp is a web-based project management tool. It is used to collaborate with project participants, plan and schedule actions, and share documentation.



**Digital Archives Migration Methodology** *Built*

### Tools



**DROID/PRONOM** *Adopted*  
<https://apps.nationalarchives.gov.uk/PRONOM/Default.aspx>  
 The National Archives UK's technical registry and file format identification tool are used for canonical identification of file formats.



**Apache Tika** *Adopted*  
<http://tika.apache.org/>



**ExifTool** *Adopted*  
<http://www.sno.phy.queensu.ca/~phil/exiftool/>  
 Metadata extraction tools such as Tika and ExifTool are used to supplement agency-provided metadata.

### Key Relationships

**State Records Act 1998**  
 S. 29 of the State Records Act allows State Records to issue guidelines about how records are to be made available to it. The guidelines apply to records in any format, including digital records.

### Key Relationships

**Disposal authorities**  
 Part 3 of the State Records Act 1998 prohibits the disposal of State records except where it is authorised. Under the Act, State Records can give permission for disposal. The usual means by which State Records permits disposal is through the approval of retention and disposal authorities.

Only digital records required as State archives under an authorised disposal authority may be transferred to the Digital State Archive.

### Access Directions

Part 6 of the State Records Act creates a framework for regulating public access to State records which have been in existence for at least 30 years (the 'open access period'). Public offices are required to make an access direction (to determine whether the records are open or closed to public access) for all their records which are in the open access period.

Agencies transferring records to the Digital State Archive must ensure that those records are covered by current access directions.

### Series control system

State Records implements the Australian Series System to describe and control the State Archives. Digital archives are linked to this system by links to agencies and series.

### Key Relationships

#### Search

<http://search.records.nsw.gov.au/search>  
 The Digital Archives team at State Records is responsible for Search, the main finding aid for the State Archives collection. Contents of the Digital State Archive are discoverable via Search.

### Tools



**Aspose** *Adopted*  
<https://www.aspose.com/>  
 A set of file format APIs for Java. Used for transformation of a number of common file formats.



**EMC Isilon** *Adopted*  
<http://www.emc.com/demand/solutions.htm>  
 EMC Isilon scale-out network-attached storage is used to store digital archives. Advantages of this system include scalability, integrity, and automated replication.



**Palfree** *Adopted*  
<https://wikilibrary.edu/id/eBay/Creator/Palfree>  
 A modified version of California Digital Library's Palfrees for Object Storage protocol is used to manage the storage of digital objects. Palfree is a filesystem hierarchy.



**MongoDB** *Adopted*  
<http://www.mongodb.net/>  
 Metadata is stored as JSON documents in the filesystem and copied into a MongoDB instance for search and reporting.



**Apache Solr** *Adopted*  
<https://lucene.apache.org/solr/>  
 Full-text is extracted wherever possible. This full-text is stored in a Solr instance for search.



**Migrate tool** *Built*  
 A workflow tool that assigns unique identifiers to digital objects, validates commitments, and moves records into the Digital State Archive.

The Migrate tool is a Java command-line application.

### Tools



**Basecamp** *Adopted*  
<https://apps.nationalarchives.gov.uk/PRONOM/Default.aspx>  
 Basecamp is a web-based project management tool. It is used to collaborate with project participants, plan and schedule actions, and share documentation.



**Receipt** *Built*  
<http://apps.nationalarchives.gov.uk/PRONOM/Default.aspx>  
 This tool provides agencies with a human and machine readable manifest of digital records successfully migrated into the Digital State Archive.

The Receipt tool is a Java command-line application.

### Key Relationships

**GA33 Source Records that have been migrated**  
 This is a disposal authority that provides for the authorised disposal of State Records that have been used as the source records for successful migration projects. It is under this authority that public offices are permitted to destroy source records post-transfer to the Digital State Archive. This authority mandates a minimum six months retention period for quality assurance purposes.

