# Information Management Advice 68 Migrating Records from Network Drives to EDRMS systems

## Introduction

*Prior to the introduction of an electronic document and records management system (EDRMS), your agency may have used a range of different systems to store digital records. These include network drives, work group shared drives and email accounts. When embarking on the implementation of an EDRMS, decisions will need to be made early in the project planning stage about the migration of existing records in these locations.*

*The key to successfully migrating documents from network drives/file shares to EDRMS is planning the project, and adequate preparation of the records prior to migration. If this process is well planned and executed properly, the migration itself is comparatively straight forward.*

The migration of records, data and information from legacy systems can be costly, time consuming and technically challenging. You should approach this issue based on business needs and risks.

Various software tools will be mentioned throughout this document, and advice on where to obtain them and any associated costs, will be included. In some cases 'scripts' will be mentioned. This refers to batch files (or similar) that your IT department can create and run for you.

## Assumptions

The information in this advice makes the following assumptions:

1) A Business Classification Scheme has already been developed in consultation with all stakeholders - Information Classification tools (2014: Advice 6)[1], Information Classification Considerations (2014: Advice 66)[2]
2) The agency has defined clear policies on use of file shares
3) IT support is available
4) Content owners have already been established - The role of an Information Asset Owner (2013: Advice 40)[3]
5) That file shares have already been structured - Structuring shared network drives for recordkeeping (2013: Advice 42)[4]

---

[1] http://www.linc.tas.gov.au/__data/assets/pdf_file/0008/334916/Information-Management-Advice-6-Information-Classification-tools.pdf
[2] http://www.linc.tas.gov.au/__data/assets/pdf_file/0018/421227/Advice-66.pdf
[3] http://www.linc.tas.gov.au/__data/assets/pdf_file/0010/398701/Advice-40.pdf

## Scoping the Migration Project

Consultation with business units is vital, without their commitment any document migration will fail. Most of the preparation work needs to be carried out by end users, under the guidance of records officers or Migration Project Team members. Some of it will need to be carried out by IT or records staff. Where possible this document identifies who should perform each task.

Steps for scoping the migration project include:

**Identify the purpose of the migration project** - is the purpose of the migration project to implement an EDRMS, or to clean up legacy documents and close off network drives?

**Identify size of issue** – create a report on the number of files and volume, break up by file type, file format and size.

**Identify the age of the issue** – how old are the documents and what are their formats? Electronic formats only have an access age/shelf life of about 5 years. If the documents are an old file type (eg MS word 2003) you will need to check that these records are still accessible, that they have not been corrupted over time, and upgrade them to a new file format in line with digital preservation requirements. *See Advice 37 Keeping Digital Records Accessible (2013)[5] and Guideline 19 Digital Preservation formats (2012)[6], for more information.*

**Identify what metadata exists about the records** - what metadata exists about the record? Will the data need to be 'cleaned up', for example retitled before migration? Is metadata mapping required pre-migration to ensure capture?

**Identify time expired records** - If the records are time expired and due for destruction; it is better to communicate this to the business area responsible. Check if there is a business reason they want them to be migrated. If not, then identify these records for disposal during the project, rather than including in the proposed migration process.

**Identify Exemptions** - During a migration project tricky document types such as databases and complex spreadsheets will be identified. These should be flagged for special consideration by technical staff as alternative management solutions may be required.

## Why migrate?

Before initiating any migration project, it is important to clearly understand what you are trying to achieve - what the business driver/s for undertaking records migration is. Reasons may include:

- Business continuity - staff have ongoing or active work in progress that need to transition to the new environment
- Preservation - as part of the agency's digital preservation strategy, identify and update file formats at risk of obsolescence

---

[4]
http://www.linc.tas.gov.au/__data/assets/pdf_file/0007/394621/Information_Managment_Advice_42_Structuring_Shared_Network_Drives_for_Recordkeeping.pdf
[5] http://www.linc.tas.gov.au/__data/assets/pdf_file/0020/390152/CENTRAL-1690921-KeepingDigitalRecordsAccessible-Advice37.pdf
[6] http://www.linc.tas.gov.au/__data/assets/pdf_file/0020/341363/TAHO_Guideline_-_19.pdf

- To support business analytics / intelligence - identification of valuable information resources, including allocation of metadata and other control tools to extract optimum value to support decision making and strategic planning
- Rationalisation of resources - identifying and reducing the number of systems, applications and disparate databases the agency ICT area is required to support, and centralising information assets into core applications
- Compliance - records have been previously unmanaged, and require classification and sentencing to enable appropriate disposal and reporting

Consider your reasons for initiating a migration project, this will help you determine what needs to be migrated (and what doesn't), and then further to identify things such as, metadata to be migrated, storage requirements, policy requirements, processes etc. These decisions will determine the scope of your project or program.

There are various reasons for migrating documents from shared network drives to EDRMS.

They include but are not limited to:

- Initial transfer of documents into a new EDRMS
- A load of documents when an organisational unit is moved into an agency
- A clean-up project to move legacy documents from file shares to EDRMS
- Bulk imports of scanned or otherwise digitised documents.
- Documents exported from a legacy system (an old EDRMS, accounting system, HR system etc.)

Regardless of your agency's reason for migrating documents, the processes are much the same and developing good processes and tools early allow you to reuse these processes down the track as new needs arise.

## What to migrate?

What needs to be migrated will be governed by the type of migration project you have - the 'why'.  If you are doing an initial load of documents from file shares, this is a large job and should not be attempted before doing a full sort of existing drives.

Shared drives will contain a massive collection of unstructured content, approximately 35-40% of this will be documents stored as legitimate shared documents (i.e. office files stored for collaboration and development). The rest will be a mixture of multimedia files (music, movie and photos), emails and email archives (.pst files), software applications, databases, backups of desktops and personal drives, 'convenience' copies of common documents. Some of this content will have arrived through legitimate process, some will not - it all has to be cleansed before any migration can take place. See Structuring shared network drives for recordkeeping (2013: Advice 42)[7] for more information.

It is important to identify what you want to take across to the EDRMS and try to avoid files that you don't need to take.  For example, most modern EDRMS systems can accommodate large complex spreadsheets but is it useful to manage these in your EDRMS. Should they be being treated as databases or applications in their own right?  Multimedia files can be managed in an EDRMS, but do they actually represent records for your agency? Are they already being managed in a business system? (e.g. a Content Management System).

---

[7]

http://www.linc.tas.gov.au/__data/assets/pdf_file/0007/394621/Information_Managment_Advice_42_Structuring_Shared_Network_Drives_for_Recordkeeping.pdf

# Key Considerations

## Metadata

It is imperative that before you embark on any migration project that you identify the required metadata for the documents once they are migrated into the EDRMS. This will include the minimum metadata according to TAHO Guidelines.

It is also essential to identify any localised metadata required for internal document and business processes. Much of your metadata may be able to be imported from the original documents but some may not. The key decision here is whether to add metadata before import, during import or manually after import.

A key issue is that the standard metadata fields in Word and similar applications are largely not compulsory unless your own agency has established business rules. So much of the data may not have been collected in the first place, and other data can be altered simply by moving documents around on file shares.

Some EDRMS may not accept imported documents that do not have a basic subset of metadata. Make sure you have researched and clearly documented what fields have to be populated to successfully load your documents into the new environment.

**6)     How do we identify what metadata we have?**

Identifying what you have can be done easily in Windows by opening an explorer window and adding or removing columns for display.
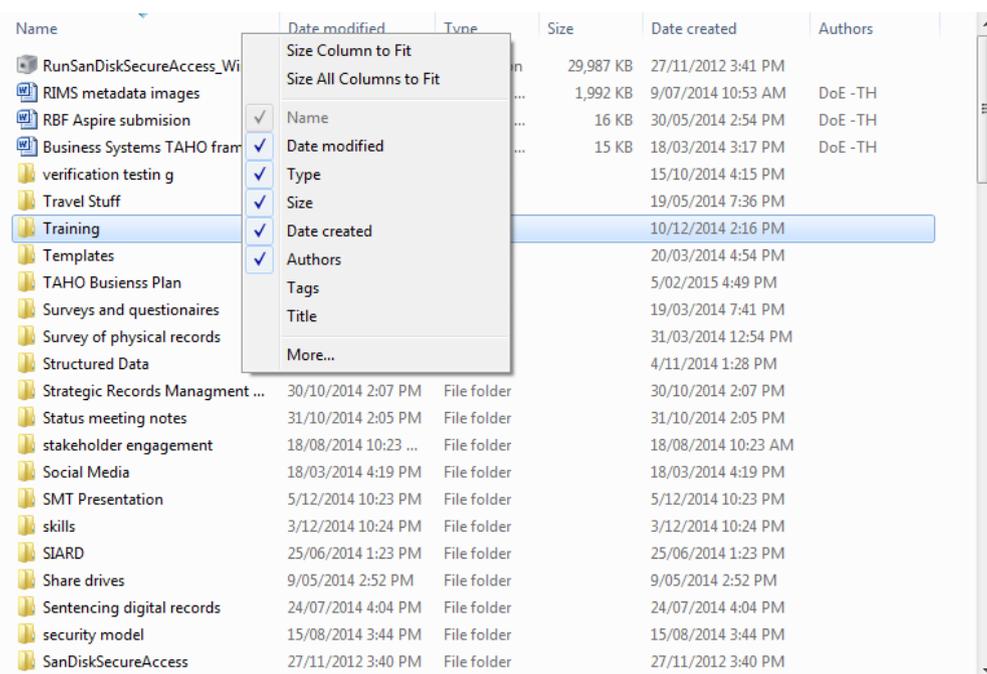


Other options that may be better suited are:

- Install a product like FreeCommander[8] which is a free file managing software product that allows you to display the contents of multiple drives. Sort by metadata and generate reports of contents on your hard drives quickly and easily. NOTE: before installing any software, contact your IT support to make

---

[8] http://www.freecommander.com/

sure it complies with agency policies and that you have the correct access/permissions to install and use the product.

- Discuss your requirements with your IT support. They may be able to generate reports providing all the information required about the documents you have, the metadata fields that are being used and file sizes etc. It is common for network management tools to have significant reporting capabilities.
- Request that your IT support or application vendors create scripts to build reports providing the specific information you need.

### 7)     How do we fix the metadata we have?

Having identified what you have currently, and compared it against the established requirements for import, you will have identified gaps or issues with existing metadata. It is recommended that you fix this *before* you attempt a migration rather than trying to do so *after*. In particular you will identify issues associated with document naming/titling. These will include but are not limited to:

- Inadequate titles, for example 0000215655555.jpg
- Names containing special characters. These can be an issue when migrating docs, as some background system processes cannot cope with them in document names.
- Poor use of case (eg. all caps, no caps etc.)

A useful tool for editing document titles is Magic File Renamer.[9] This is free to download and install, and cheap to license if you need to do more than a couple of updates. It allows the building of 'replace' lists to remove special characters, adding folder information to titles, correcting case issues (i.e. converting to lower or upper case, or selecting sentence case).  As mentioned before, talk to your IT support before attempting to install.

## Versions

Particular consideration needs to be given to versions of documents.

It is sometimes necessary to import multiple versions for long term management and support of process, compliance, digital preservation of permanent value records, or other reasons. For example:

- If previous versions have been published and used for decision making.
- If a document is still under development/review

Most EDRMS will not easily support the importation of multiple versions of the same document through their existing import functions. This does not necessarily mean it can't be done, simply that it may be difficult and therefore not cost effective to do.

It may be possible to flag multiple versions and use a script to import them as versions of a single document.

The most straight forward method of import may be to import one version only, either the current or the first depending on local needs, and manually import any additional versions. The choice between first and current will depend on the processes employed, and the reason for import. In most cases whichever version you import will be treated as the original or first version. It is strongly recommended that metadata is added to the imported document to identify what choices were made regarding versions and reasoning for the import.

Before deciding which method of import to investigate or use, you need to look at how many versioned documents you have and how clearly they are currently titled and recorded.

---

[9] http://www.finebytes.com/mfr/

## Illegal Characters

Illegal characters refer to any characters in titles that may cause issues with handling documents during the migration process. It refers to characters that are not accepted by the base operating systems. They will include but are not limited to:

\ / : * ? " < > |

It is strongly advised that you discuss this list with your IT support, and that you coordinate the removal of these from document titling using options discussed in the previous section - 'How Do we Fix the Metadata We Have'.

## Training

Training is an essential part of any migration effort. Key items that will need to be covered include:

- the BCS (business classification scheme) that is to be used throughout the migration, and in the new system.
- Impact: staff need to know and understand how the migration will affect them, what role they will play in it, and what the outcomes will be.
- How to use the new software including the EDRMS you will be using (e.g. TRIM)

Training needs to be introduced from the very start of engagement with each business unit, it can be informal in the form of staff meetings about the move or more formal. Pay attention to your own organisation and how its units work to get the best return on training you can. Refer to TAHO Advices on Change Management (Advice 55 and Advice 20)[i] for guidance on communications planning, change management initiatives and staff training strategies.

## Talking to IT

The relationship you have with your IT department and application vendors is pivotal to the success of any migration project. It is as important for you to be aware of the needs of your IT department as it is for them to be aware of yours.

Be particularly careful of language when talking to IT people. For instance the term 'file' has different meanings in IT and in Recordkeeping although the roots are the same. It can be either a container or the thing contained in the container.

Also be mindful of phrases being used, a common example is that something 'can't be done'. This can often mean:

- 'we don't have the skills to do that'
- 'we have never done that before'
- 'we don't know what you are talking about'

It can also mean:

- 'we did not allow for that in our quote'
- 'we don't have the resources'

These comments should always be questioned. The reality is you need certain things done, you need to know how, and at what cost they can be achieved. If you do not get a reasonable answer, make sure your question is escalated to someone who can give you a clear itemised explanation of the position.

Remember that in most cases a document migration project brings gains for the IT department, as well as for Records departments. Focus on these benefits when dealing with your IT support. Gains include but are not limited to:

- Clearing space on network drives
- Identifying and managing unwanted content on drives
- Improved performance
- Compliance with agency IT policies
- Compliance with agency Record Keeping policies

## Business Process Analysis

Process analysis should be undertaken for all records that are stored on your network drives that are a part of a business process, and will be included in your migration. This helps to verify that you are taking the correct documents and applying the correct metadata.  It will help identify any potential issues and also any areas where you can create wins for your agency.

## Automated Sorting

It is possible to do some automated sorting of shared drives.  Manual sorting is the only effective way to properly sort, but the reality is that shared drives can contain an extensive history of random document and folder creation, and agencies rarely have the resources available to undertake a purely manual sort.

Some suggested options to assist in preparation of documents prior to migration are:

- Eliminate linked files. Many EDRMS will struggle with importing linked files. Often links will be broken on importation as the links rely on document location in file paths for their addressing.
- Date sort. Sorting files into a new file structure to limit the number to deal with can be very helpful. For example, if you have decided to only import current or active documents, then sorting only the last 6 months of documents into a new structure greatly reduces the amount of effort required.
- Sort by document type. Once you have decided what document types you are importing to the system, identifying and sorting these in groups can improve the efficiency of a sort process.
- Flatten file structure. Often with shared drives, child folders can be created very deep within the structure leading to too many levels which makes sorting very clumsy. By flattening a folder structure this can prevent hidden folders being missed.

## Planning your Migration Project

Once you have worked through all of the considerations and made some decisions about the migration project scope you will need to document this in a project plan.  This plan should be based on a sound project management methodology. Refer the Tasmanian Government Project Management Guidelines if your agency does not have a preferred formal methodology.  A part of this plan should include documenting and communicating the steps in the project for your stakeholders.  Attachment 1 to this Advice provides an example.

## What to do if you are not going to migrate?

Your agency will need to properly support and maintain information and records from superseded systems that are not going to be migrated to ensure ongoing access for as long as required for business and legal purposes.

It is also important to have appropriate procedures in place regarding future use of superseded systems. An example may be disabling network drives so that all future corporate records **must** be saved to the EDRMS. However, 'read only' or 'copy only' access may be provided to end users so they can continue to access historic information they require. You need to choose an approach that allows some flexibility, but also meets the business objectives of the EDRMS implementation.

## Glossary

**Script** – a small non-compiled computer program written to automate specific task on computers and computer systems

**Metadata** – data about data.  In records management, recordkeeping metadata is data that describes the context, content, and structure of records.

**EDRMS** – Electronic Document & Records Management System. A type of content management system that refers to the combined technologies of document management and records management systems, as an integrated system.

**CMS** – Content Management System. The function and use of content management systems is to store and organise files, and provide version-controlled access to data. Enterprise systems offer functions including web-based publishing, format management, indexing, search and retrieval.

**Multimedia** – Generic term meaning any or all of photo, movie and sound files

**File Share** – shared computer drives which have access granted to multiple users

**Network Drive** – computer hard drive that is accessible from a computer network rather than just the local terminal

# Further Advice

For more detailed advice, please contact:

Government Information Strategy Unit
Tasmanian Archive and Heritage Office
91 Murray Street
HOBART TASMANIA 7000
Telephone: 03 6165 5581
Email: gisu@education.tas.gov.au

# Acknowledgements

- Implementing an EDRMS: Key Considerations, National Archives of Australia, 2011
- Tasmanian Government Agency EDRMS Implementation Projects

## Information Security Classification

This document has been security classified using the Tasmanian Government Information Security classification standard as PUBLIC and will be managed according to the requirements of the Tasmanian Government Information Security Policy.

## Document Development History
## Build Status

| Version | Date | Author | Reason | Sections |
|---------|------|--------|--------|----------|
| 1.0 | April  2015 | William Bester | Initial Release | All |

## Amendments in this Release

| Section Title | Section Number | Amendment Summary |
|---------------|----------------|-------------------|
| | | This is the first release of this document. |

## Issued: August 2015

## Ross Latham
State Archivist

# Attachment 1

The following migration process was used in a large agency that was implementing an EDRMS for the first time. The agency had at least 20 years worth of digital records stored on network drives. The scope of the migration project was to identify the previous 6 months worth of digital records and classify them according to the agency BCS ready for migration into the system. A separate project was scoped to manage the legacy data. The following folder structure was established on Business Units network drives to support automated and manual sort processes.

**For Deletion** - for multiple copies of the same records and records that the business unit no longer required. These records were to be reviewed, classified and destroyed, when time expired following the migration.

**For Archiving** - for Business Unit records that were no longer being actively used and were not required for the day to day running of the business unit. These records were to be reviewed, classified and destroyed, when time expired or migrated to the EDRMS following the initial EDRMS implementation.

**For Migration** - current business unit records required for the day to day running of the business area as identified by the business unit or move to this folder by automatically selecting the previous 6 months worth of records created on the network drive.

**Non Migration Docs** - This folder is for any records which are in formats that are not easily managed within an EDRMS. For example databases, CAD drawings or spreadsheets with links where the business unit does not want to break the links by moving them to the EDRMS.

**File Plan** - this file is created to hold a copy of the BCS from the EDRMS. Business Unit staff then review the content of the 'For Migration' folder and move it into the copies of the files from the BCS ready to be migrated.

| Action | Aim | Responsibility |
|---|---|---|
| Create 'For Deletion' 'For Archiving' 'For Migration' 'Non Migration Docs' 'File Plan' folders | Provide an areas within which historic / unwanted documents may be stored. Delineating between current and non-current documents and migrating and non migrating documents.<br><br>The folder creation process can be easily done with a script. Discuss with your IT support | Business Unit |
| Business Unit decide on a date for the automated sort | Assists the business unit to sort documents to migrate | Business Unit |
| Preparation for the automated sort process | Runs reports to identify duplicate documents, document names with illegal characters (*,./), Identify documents with links and then sort documents by date.<br><br>This can be done with scripts and in some cases with disc management tools, discuss with your IT support | Support Technician |

| | | |
|---|---|---|
| Run automated drive sort process. | This process copies the entire File share structure into both the 'For Archiving' and 'For Migration' folders. Documents on the File Share are dispersed between the 'For Migration' and 'For Archiving' folder based on a 'last accessed' date, as specified by the business unit, and EDRMS compatible file types.<br><br>Again scripts or disc management tools are used | Migration Team / Business Unit |
| Manual Sort of 'For Migration' and 'For Archiving' | Ensuring that all current business documents are 'caught' within the 'For Migration' folder | Business Unit |
| Conduct BCS training and introduce business unit to the EDRMS system. Explaining how the BCS and EDRMS work together. | Impart an understanding of the BCS and how it works in conjunction with the EDRMS system | Migration Team |
| Run flattening of 'For Migration' process. | Ensures that all documents are stored in folders at a single level. Facilitating the classification of all documents stored within the 'For Migration' folder.<br><br>This is a job for scripts again | Migration Team / Business Unit |
| Classify all folders within the 'For Migration' folder using the BCS. Folders are to be classified to the third level. (Recommended, but this depends on the classification depth you have decided for your implementation)<br><br>(Incorporating a feedback mechanism for the business unit to comment on suggested classification.) | Provide a functional classification for all current business unit documents. | Migration Team / Business Unit |
| Create new files identified as part of the classification process in EDRMS. | Generating file numbers that can then be incorporated in the folder names within the File Share drive file plan | Records Team |
| Move documents, from the flattened 'For Migration' structure to the appropriate location within the File Share file plan as per the classification suggested by Migration team and accepted by the business unit | Construct a file plan very similar to the file plan structure evident in Hummingbird.<br><br>Option: This process can be done at classification time | Business Unit |

| | | |
|---|---|---|
| Rename folders containing documents within the File Plan incorporating the file number, as found in the EDRMS, into the folder name.<br>Use the following convention:<br><folder name>---<file number><br>e.g.<br>Records and information management project plan---CEN_567/1 | Applying file numbers to the 'Bottom level' folders within the File Share file plan allows the automatic load process to match files to folders. | Business Unit |
| Run Share drive 'File Plan' report and check. | Once all documents have been moved to the Share drive 'File Plan' a check of the document names must be performed to ensure that:<br><br>No illegal characters may be found in documents names.<br><br>Document names adhere to the document naming convention. | Migration Team |
| Provide file plan report feedback to the business unit. | Allows the business unit to correct any problems with the Share drive 'File Plan' prior to the document load process | Business Unit |
| [If there is a delay between completion of the classification process and installation of EDRMS]<br>Run the Share drive file plan report one day prior to the automated document load process. | Allows the Migration team technical officer to identify anything within the file plan that may lead to load failure.<br><br>Double check file numbers within the file plan to ensure that file number transposition hasn't occurred. | Migration Team |
| Forward details of any potential issues arising from the Share drive File Plan report to the business unit for correction | Allows the business unit to address issues prior to the document load. | Business Unit |
| EDRMS software training | | Migration Team |
| EDRMS Implementation Day | | All |

[i] Managing Change in EDRMS implementation projects(2014: Advice 20)
Change Management - Preparing for Change (2014: Advice 55)