

Managing digital records: 4

Contents

4.	Effectively manage the migration of your digital records.....	1
4.1	Recognise that migration is a high risk process	1
4.2	Before migration, be aware of key record requirements.....	1
4.3	Use migration to preserve the long term stability of key business records	8
4.4	Be aware of triggers for migration and different migration types	9
4.5	Plan for migration	10
4.6	What about contractors?	18
4.7	Perform pre migration testing	19
4.8	Perform the migration	19
4.9	Perform post migration testing.....	20
4.10	Make records of your migration.....	21
4.11	Keep source records for at least six months.....	21
4.12	Think strategically about migration.....	23
4.13	Checklist for migration.....	23

4. Effectively manage the migration of your digital records

4.1 Recognise that migration is a high risk process

Migration is a preservation activity that transfers records from one hardware or software configuration to another, or from one generation of technology to another.

Migration is necessary because the many protocols and software components that enable records to be read and used are constantly evolving. Their evolution is rapid and compatibility with earlier versions is often not retained, especially over periods longer than a few years. Without migration, access to important organisational records would be lost.

Migration however is a high risk process. It changes data and therefore threatens the authenticity, integrity and even the existence of records.

This section does not provide technical detail about migration operations, such as choosing new data formats or developing suitable migration methods. Generally organisations will have good technical and vendor support to provide guidance on these issues. Instead this section focuses on recordkeeping issues and provides information to help mitigate the risks associated with migration and to protect record authenticity, integrity and useability.

4.2 Before migration, be aware of key record requirements

Records have certain defining features that must be supported during migration operations. Understanding these features is critical to maintaining record authenticity, integrity, reliability and useability during migrations. The defining features that you need to understand before you plan for and implement migration operations are:

- records are complex
- metadata is critical
- essential characteristics must be preserved

Records are complex

Records are not simply data. In order to serve as evidence and information they are comprised of a complex of related information:

- structure – the form and layout of the record
- content – the informational value of the record – this could be simple text or a complex aggregation, such as a word processed document containing a spreadsheet or a web page containing a variety of images
- context – information about who created the record, why they created it, how it has been managed and what other records it is related to.

All of this information must be maintained during migration to preserve the evidential and informational value of records.

Metadata is critical

Metadata is used to describe and manage records. It is generally the means by which much of a record's context is documented and is the ultimate means by which the integrity and authenticity of a record can be proven. It is therefore critical that it is preserved and that connections between a record and its metadata are maintained during migration.

Metadata is also critical to system functionality. In any migration process it is vital to consider the following requirements. Failure to do so will jeopardise the integrity and useability of organisational records.

Requirement:	Why:
Perform a metadata mapping between the original and the target system	This will ensure that all necessary metadata fields and their values are preserved following migration.
Migrate records management controls	Records management controls are disposal authorities, security classifications and record classification tools. Much functionality may be lost and much expense incurred if these tools and the functionality they deliver is not adequately identified and migrated to the target system. It is important to be aware that complex relationships can exist between one or more of these tools and some or all records in the system. These relationships must be preserved during migration.
Maintain any additional functionality driven by metadata	If metadata is used to automate activities (such as disposal or preservation actions) or if metadata reuse or other forms of automation have been used in the system, then this functionality must be safeguarded during the migration.
Identify the different varieties and aggregations of records in the system and	The records in your system may be individual documents, aggregations of related documents (files or volumes) or aggregations of files

interdependencies that exist between them	<p>(series). These aggregations and the implicit relationship connections that exist between them must be safeguarded during the migration as these relationships are part of the evidentiary context of the records.</p> <p>Aggregations can also be used to apply management rules. For example, if they are uniform, disposal or other rules can be specified at the file or series level and a business rule used to identify that these rules should apply to all records nested within the file or series container.</p> <p>In such situations, this functionality must be replicated in the target system or the implicit metadata represented by the aggregations must be made explicit in the target system.</p>
---	---

Essential characteristics must be preserved

Essential characteristics are those features that are critical to a record's meaning, use or organisational value. All migration operations should be designed to preserve the essential characteristics of the records being migrated.

Each organisation needs to determine the essential characteristics that apply to their own specific records. Generally, essential characteristics will differ according to record type and the business purpose served by the record.

Examples of identifying essential characteristics:

A report contains a map where colours are used to signify different agricultural areas. These colours have meaning: the report could not be interpreted accurately if these colours were not preserved. Therefore for this report the colours are an essential characteristic and any migration performed on this report must ensure that this essential characteristic is maintained. However if the colour of a record is a slight tint to the background of a letter, then this is not likely to be an essential characteristic and is not necessary to preserve in migration operations.

When migrating email, you will generally not regard the appearance of messages as essential. With some text records, however, you may regard their appearance as an essential characteristic as it adds meaning. The appearance of a table in a Word document is an example. If this is altered, the information presented in the table can lose its meaning. This then is an essential characteristic that must be preserved.

Recent research in the Netherlands sought to identify the essential characteristics of four common types of business records – word processed, email messages, spreadsheets and databases.¹

The essential characteristics identified were based on what the project saw as the five basic attributes of digital records:

- content
- context

¹ The full research produced by the Digital Preservation Testbed Project can be accessed from the Netherlands' *Digital longevity website*, viewed June 2008, <<http://www.digitaleduurzaamheid.nl/index.cfm?paginakeuze=185&lang=en>>

- structure
- appearance, and
- behaviour.

The following tables identify aspects of content, context, structure, appearance and behaviour that the Testbed research identified as essential to preserve for text, email, spreadsheet and database records.

The essential characteristics listed in the following tables may be useful for your own preservation planning. They may provide a guide for determining the essential characteristics of text, email, spreadsheets, databases and other types of records created by your organisation that must be preserved.

It should be noted that the essential characteristics that follow are generic and are based on record type. There may be other specific characteristics unique to your business that you will need to identify and manage in order to ensure that all the essential characteristics of your own records are preserved.

Word processed documents

<i>To preserve:</i>	<i>You need to maintain:</i>
Context	<ul style="list-style-type: none"> • metadata identifying: <ul style="list-style-type: none"> • the name of the creator and creating organisation • the business process that produced the record • its date of creation • its specified relationships to other records • its original and current file formats • a history of the recordkeeping actions performed on the record.
Content	<ul style="list-style-type: none"> • all content – text, images or otherwise. This includes: <ul style="list-style-type: none"> • page numbers • headers and footers • automatically created content such as tables of contents and indexes • document properties • comments that may be attached.
Structure	<ul style="list-style-type: none"> • the logical relationships that exist between the various components of the text document when rendered onscreen.
Appearance	<ul style="list-style-type: none"> • features of the record appearance which convey meaning. These may include: <ul style="list-style-type: none"> • bolding • italics • underlining • font size/style. <p>Other aspects of the appearance of the migrated record may change from the appearance of the original provided that the</p>

	meaning of the record is left unchanged.
Behaviour	<ul style="list-style-type: none"> any functionality in the record that ensures it is understood such as the ability to view embedded graphics behaviour that is counter to ongoing integrity of the record such as an automatically updating date field should be 'turned off'.

Email records

<i>To preserve:</i>	<i>You need to maintain:</i>
Context	<ul style="list-style-type: none"> essential header elements: <ul style="list-style-type: none"> the email address, the organisation and the full name of the sender the email address, the organisation and the full name of all recipients for outgoing messages, the date and time the message was sent for incoming messages, the date and time the message was received, as well as the date and time the message was sent the subject of the message security and/or confidentiality settings the file name and the file format of any attachments data about original and current file formats a description of the business process that produced the record any specified relationships to other records a history of the recordkeeping actions performed on the record.
Content	<ul style="list-style-type: none"> all content – this includes: <ul style="list-style-type: none"> extra files such as photos or images inserted in the original message links and hyperlinks.
Structure	<ul style="list-style-type: none"> the structure of the original message attachments and inserted items.
Appearance	<ul style="list-style-type: none"> the meaning of the message, but the appearance of the message does not need to remain the same.
Behaviour	<ul style="list-style-type: none"> the ability to open and access attachments links to other documents, including the title of the document.

Spreadsheets

<i>To preserve:</i>	<i>You need to maintain:</i>
Context	<ul style="list-style-type: none"> • metadata identifying: <ul style="list-style-type: none"> • the name of the creator and creating organisation • the business process that produced the record • its date of creation • its specified relationships to other records • its original and current file formats • a history of the recordkeeping actions performed on the record.
Content	<ul style="list-style-type: none"> • the content of the spreadsheet, including the content of any possible charts and other inserted objects. This includes: <ul style="list-style-type: none"> • both the standard display of the worksheets (the entered values and the results of calculations using formulae) and the underlying layer containing formulae and inserted functions.
Structure	<ul style="list-style-type: none"> • worksheets in the correct sequence with the correct names • the existing row and column structure of the document which provides each cell with its unique designation – ie migration must not change the coordinates of a cell from A3 to B3 as this could cause the formulae embedded in the document to produce a result that differs from the original value • inserted charts or objects • links to related spreadsheets. Related spreadsheets must also be preserved if these provide values to cells • links between calculations and the formulae that produce them.
Appearance	<ul style="list-style-type: none"> • the appearance of inserted objects. <p>The appearance of the spreadsheet and charts may differ from the originals.</p>
Behaviour	<ul style="list-style-type: none"> • the ability of formulae to carry out a calculation or recalculation. If records are no longer actively used for business, this functionality will need to be disabled so that the record is not constantly altered each time it is opened. <p>It should be noted that European testing showed that newer versions of spreadsheet applications sometimes use mathematical formulae in a manner different to the older version, thereby yielding a different result. The ability to apply formulae should therefore usually be disabled when the record is captured into a recordkeeping system. Before migration, it should be checked that all formulae are actually disabled.</p> <p>Remember, the preservation of an authentic record requires the retention of its content of the record at the time it played a</p>

	role in the business process. This is particularly important in relation to spreadsheets which are easily subject to change.
--	--

Databases

<i>To preserve:</i>	<i>You need to maintain:</i>
Context	<ul style="list-style-type: none"> • metadata identifying: <ul style="list-style-type: none"> • the name of the creator and creating organisation • the business process that produced the record • its date of creation • its specified relationships to other records • its original and current file formats • the name and version of the query languages that are used • a history of the recordkeeping actions performed on the record.
Content	<ul style="list-style-type: none"> • the actual content of the tables • the queries used so that the required content can be represented.
Structure	<ul style="list-style-type: none"> • the physical structure of the database. This includes: <ul style="list-style-type: none"> • the tables • relationships between the tables, including the constraints • the views • field attributes • the structural composition of the data as presented onscreen. <p>The logical structure of the database can be preserved in an entity relationship diagram or an XML schema.</p>
Appearance	<ul style="list-style-type: none"> • the onscreen representation.
Behaviour	<ul style="list-style-type: none"> • the behaviour of the user application <ul style="list-style-type: none"> • this can be preserved in the form of descriptions of system-supported functions, as well as screenshots of the displays used for entering and amending data, generating reports, etc. It can also be preserved in system documentation and user manuals.

Essential characteristics can come from business requirements and these too must be maintained during migration. The need to maintain essential characteristics can sometimes limit your migration choices. For example, a recent experiment in the United States sought to determine whether complex three-dimensional, geometric CAD (Computer Aided Design) records of high tolerance machined piece parts could be migrated from their native CAD environment to an open source archival format to facilitate their long term preservation. The

experiment showed that, at present, it was not possible to successfully migrate all aspects of the records to the non proprietary format. The open format could not adequately represent the fine accuracy and measurement levels (down to a millionth of an inch) that were necessary to sustain the accuracy of the engineering drawing and consequently the accuracy of any product subsequently created from that drawing. Because the exacting engineering requirements that are an essential characteristic of these records could not be reproduced, migration and maintenance options for these records are limited to the proprietary format in which the records were created. ²

4.3 Use migration to preserve the long term stability of key business records

In every organisation there will be some business records that will be required to be retained for legal and informational reasons for periods of more than 30 years. Some records have extremely long retention periods, for example some need to be retained for the life of an asset, such as a building or a bridge.

Long term stability is important, particularly for records with high long term value. Migration however is an inherently risky process and repeated migrations can be particularly damaging to records.

Migrating records to a stable long term storage format is a means to achieve long term stability and to protect high value records. A stable long term format is a widely-used, non-proprietary, platform-independent format that is either uncompressed or lossless with, where possible, freely available specifications. These formats tend to have a long expected life which helps to minimise the number of migrations that a record will be subject to. See [1 Make digital recordkeeping achievable for your organisation](#) for a list of the stable long term formats recommended for records of long term value.

Migrating records to stable long term formats can also save time and money because this approach minimises the number of migrations that you will potentially be required to perform.

You should consider which of your organisational records could benefit from being migrated to a stable long term storage format. In general, records are migrated to long term storage formats when they are no longer required for active business operations. Records can however be migrated to standard storage formats earlier in their lifespan if this meets the needs of your organisation. For example, one large organisation migrates all its client files which have very long retention periods and long term rates of business use to PDF five years after they are created.

You can perform the migration to stable long term formats yourself or you can use freely available tools such as Xena. Xena is free and open source software to aid in the long term preservation of digital records. It can be used to convert records to an XML-based archival data format (a process known as 'normalisation'). Xena is an acronym meaning 'Xml Electronic Normalising for Archives'. Xena was developed by the National Archives of Australia.

The Xena software can be used to:

- detect the file formats of digital objects

² This case study is discussed in L Duranti and R Preston, *International research on permanent authentic records in electronic systems (InterPares) 2: Experiential, interactive and dynamic records*, 2008, pp.34-35, viewed June 2008, <http://www.interpares.org/display_file.cfm?doc=ip2_book_complete.pdf>

- convert digital objects into open formats for preservation (known as 'normalisation').

Native formats that XENA can currently convert include:

- MS-Word, Excel, Powerpoint and Project
- OpenOffice.org Writer, Calc, and Impress
- RTF
- PST email format
- TRIM email format
- MBOX email format
- Comma Separated Files (CSV)
- JPG, GIF, TIFF, PNG, BMP, PCX
- HTML
- Plaintext (various encodings)
- PDF documents, and
- XML.

Further assistance:

For more information on Xena or to download the application see <<http://xena.sourceforge.net/links.php>>.

When migrating records to a long term storage format it is critical that the migration follows the planning, testing and documentation requirements outlined in the remainder of this guideline. Failure to do so will jeopardise the integrity of the records and the success of the migration.

4.4 Be aware of triggers for migration and different migration types

The different triggers for migration and the different types of migrations performed in response to them include:

<i>Migration from:</i>	<i>Triggered by:</i>
<ul style="list-style-type: none"> • one internal organisational business system to another 	<ul style="list-style-type: none"> • changing business needs that lead to the adoption of a new business system • technological changes that require the update of business systems • the need to transfer records from active business systems to long term storage systems • other internal and external drivers.
<ul style="list-style-type: none"> • one organisation to another 	<ul style="list-style-type: none"> • administrative change that requires the transfer of functions or activities from one organisation to another • the adoption of shared service

	<p>arrangements that require the transfer of records from one organisation to another</p> <ul style="list-style-type: none"> • other internal and external drivers.
<ul style="list-style-type: none"> • one organisation to State Records 	<ul style="list-style-type: none"> • requirements for the management of State archives, where State Records is required to take control and custody of State archives • other business drivers.

The same broad requirements generally apply to all forms of migration and this should be acknowledged in your organisation's rules on migration.

There are effectively three different levels of migrations:

- minor functionality change – records are moving from one system to another that is very similar
- medium levels of functionality change – records are moving from one system to another that offers different, usually upgraded functionality, and
- migration to a complete new system.

The same requirements, performed to different degrees, apply to all levels of migrations.

<i>Migration type</i>	<i>Tasks that need to be performed</i>
Minor functionality change	Project planning, checking of provided export/import gateways that will perform migration, pre and post migration testing
Medium levels of functionality change	Project planning, checking of provided export/import gateways that will (usually) perform migration, pre and post migration testing, training and improved user documentation
Migration to a complete new system	Extensive planning, (potentially) writing of special purpose software to move the records and the application functionality into a new technological environment, pre and post migration testing, significant change management issues

4.5 Plan for migration

Once you know you need to perform a migration, there are numerous issues you need to consider when developing your plan of how your migration will be performed. The following tables identify some of the data and system issues you may need to consider.

Data issues to consider in migration planning

<i>Issue:</i>	<i>Discussion:</i>
---------------	--------------------

Data cleaning	<ul style="list-style-type: none"> • What is the quality of back end tables? Are the staffing details in these tables up to date? Are people who have left the organisation still recorded there and listed as still holding records? You do not want to migrate this data. You should clean it up first. • Does a file audit need to be conducted to identify current locations of paper-based records? • Is the latest version of retention and disposal authorities referenced – for example GA28 and not GDA2 for administrative records? If not, do you want to rectify this prior to migration?
Making decisions about metadata	<ul style="list-style-type: none"> • Sometimes people decide not to migrate all metadata components of a record. This must only be the case for those metadata elements that have no ongoing business or accountability relevance for the organisation. If for technical, capacity or management reasons you decide not to migrate all metadata elements, these decisions should be fully considered and comprehensively documented.
General records clean up	<ul style="list-style-type: none"> • One large public sector organisation held 'Records Week' in the week preceding their records system migration. Staff were required to clean out shared directories and email systems, file all records and return hard copy files to the registry. This helped to ensure that as many records as possible were in the system prior to migration.
Implement disposal	<ul style="list-style-type: none"> • Can certain records be disposed of rather than migrated? It is not efficient to migrate records that can be disposed of. • Migration is a good opportunity to see what records you have stored in secondary storage areas and to cut down on these records if possible.
Identify rogue and independent systems	<ul style="list-style-type: none"> • Staff frequently maintain personal systems that they use to control business records. Migration can be used as a trigger to bring these rogue or independent systems within broader business systems and to gain greater control and accessibility to all business records of the organisation.
Format requirements	<ul style="list-style-type: none"> • Ensure you understand the hardware, software and format requirements of all the records requiring migration.
Non standard records	<ul style="list-style-type: none"> • Be aware of any non standard records that you will need to migrate and ensure these are incorporated in your plans.
Older records that have	<ul style="list-style-type: none"> • Check the earliest records in your system.

undergone multiple migrations	Look at their quality, check their dependencies, look at what is going on with these records and what is required to support their ongoing maintenance.
Records stored in diverse environments and different formats	<ul style="list-style-type: none"> • Ensure that all records requiring migration are actually migrated. Individual records should not be inadvertently left out of the migration action. You should ensure that all relevant online and offline records, and those at secondary storage or data warehousing facilities are included where appropriate.
Data cleansing (or data massaging or scrubbing)	<ul style="list-style-type: none"> • Data may need cleansing or massaging to get it from an older system to a newer one. • Cleansing requirements may be format related – ie some older formats may need formatting removed and poorly written code cleaned up, usually through a vendor approved or custom script. • Cleansing requirements can also be system related – older systems tend to store large records as separate pages while newer systems tend to save them as one object. You may choose to clean these older records by converting them to a single document before migration, or you may leave them and build transformation rules into the target system. • Any data cleansing needs to be fully documented to help prove the integrity of the migrated records.³
Consolidating records	<ul style="list-style-type: none"> • One very large organisation chose not to migrate all its records. For example, its old system contained 150 000 records of destroyed files. These were translated into one file and this was migrated and maintained, rather than each of the individual records. Any change of this type needs to be well documented to help maintain integrity.
Movement history data	<ul style="list-style-type: none"> • Importing movement history as active metadata is very difficult as it involves migrating and maintaining a 20 year history of staff. At migration it may be easier to maintain movement histories as a record and start movement tracking afresh in the new system. Again, any such changes need to be documented. This approach may be particularly useful if you have not conducted a records audit prior to migration to identify missing records or records still marked out to

³ These points were principally drawn from Part II, Section 1.2 of ARMA International, ANSI/ARMA 16-2007, *The digital records conversion process: Program planning, requirements, procedures*. No online reference.

	staff that have left the organisation.
Create explicit metadata fields where necessary	<ul style="list-style-type: none"> • Some systems do not store sender, receiver and date received metadata from email messages in separate fields. You need to establish this yourself if required. Searching and accessibility can be badly limited if this functionality is not enabled. • Some metadata is not explicit. For example, some disposal metadata is applied through inheritance from a parent record. At migration, all metadata values should be made explicit or the functionality preserved to allow these values to be inherited from a parent record.
Talk to users	<ul style="list-style-type: none"> • Talk to users of the system to make sure that you are planning to keep all the information they need to keep using the records.
Decompression and unencryption	<ul style="list-style-type: none"> • Determine whether any records need to be decompressed or unencrypted before they can be migrated.
Digital signatures	<ul style="list-style-type: none"> • Identify whether any digital signatures need to be migrated and how this will be achieved.
The quality of encoding schemes or 'picklists'	<ul style="list-style-type: none"> • In many systems, picklists are used to populate a significant proportion of your data. When moving to a new system you need to review the quality of your existing picklists and make sure that they still reflect current requirements. Check that the available values are actually the values that bring the most benefit. Your new system may have increased functionality and may have the capacity for new or expanded picklists so make sure you investigate this possibility. • If you alter or remove a picklist that was present in your old system, make a record of what it was and document any changes in a report or file note.
User understanding of the purpose of data fields	<ul style="list-style-type: none"> • Look at how people are using existing data elements. Are people using specific fields in a variety of different ways? Does your title element contain disposal, location and agent information, as well as title information? This is a common problem that really limits system usability. • You can fix this problem by issuing very clear procedures with your new system that explain exactly how the system should be used. Then you need to monitor system useage and follow up with those who use the system inappropriately.
Multiple values in single metadata fields	<ul style="list-style-type: none"> • It is important to note that fields such as the title field described above that contain a variety of different information elements can cause problems during the migration. Migration is generally a one to one mapping of a data element in the old system to a data element in the new system. In some situations it may not be possible

	<p>to map multiple values in one element to various different elements in your target system.</p> <ul style="list-style-type: none"> If you encounter a situation like this, you may need to look at manually adding data to some fields to compensate, migrating some data to a Notes field or undertaking a more complex migration than you originally planned.
<p>Maintaining countdowns that are already in operation</p>	<ul style="list-style-type: none"> If your current system has implemented disposal triggers and has already started calculating retention periods based on these, you must make sure that these calculations are maintained unchanged in your migration. For example, if you have a group of records that are to be kept for five years after audit and the audit was conducted two years ago and your system has calculated the destruction due date based on these events, then the system countdown must be maintained accurately and not start again following migration. This is a common problem at migration that can result in a significant amount of resentencing work, so make sure it does not become a problem during your migration project.
<p>Getting an accurate idea of the range of metadata in your system and how it is actually used</p>	<ul style="list-style-type: none"> If your system is large and if it has been in operation for a long time it is likely that it has been used in different ways by different people over time. You should run a variety of different reports to enable you to get a comprehensive overview of all the data in the system and the different ways it has been recorded. This will help you to ensure that your migration plans are comprehensive and accurate and will help you to avoid any potential data loss in migration. As an example, the migration planning team in one large organisation assumed that all business units used the Title field to document record title information, but ran reports and generated extracts to confirm their assumptions. In the process they discovered that one business unit had been using a minor descriptive field to record title information and this descriptive field was not scheduled for migration. The team subsequently changed their migration plans to ensure that this field was migrated to the new system. Without their checks, necessary data would have been lost from the system.
<p>Checking the comprehensiveness of your Date information</p>	<ul style="list-style-type: none"> Dates are very important in recordkeeping. Records systems generally have the capacity to document a lot of dates about record creation, registration, disposal, access or forthcoming recordkeeping events. New systems have greater capacities for date application than old systems. In addition, applying date metadata was not done as rigorously as it should have been in some old systems. Both of these scenarios can lead

	<p>to problems in migration.</p> <ul style="list-style-type: none"> • New systems can often have a default date setting as standard out-of-the-box functionality. If, for example, you did not add 'creation date' to all your files in your old system and if 'creation date' is a mandatory field in your new system, at migration the system may automatically apply the current date to all blank date fields, irrespective of when the file was actually created. This can cause problems with accurate searching, disposal, legal requests etc. • To prevent these problems you need to check the comprehensiveness of your existing date metadata and the business rules that apply to dates in your new system. This will enable you to anticipate any problems and take steps to resolve them. For example, you may decide to perform some data cleansing prior to migration by back capturing accurate date information. This is a more cost effective approach than allowing inaccurate data to be recorded in the system. • In most systems you can run a report to identify any blank fields. You should do this to identify whether there is any data that is missing or that will be automatically rewritten by the new system at migration.
<p>Records that are currently stored in different databases that will be consolidated in a single database following migration</p>	<ul style="list-style-type: none"> • One organisation was using its migration to consolidate records stored across different databases into one database. In the multiple database structure, records in each database were given a prefix based on the name of the section that owned the database. For example, the Finance and Administration records all had identifiers beginning with 'F', ie F07/1234 etc. Migration planning staff realised that the records of the Property division and the Planning division had the same prefix and consequently all had the same record identifiers. At migration they ran a business rule to change all the identifiers in the Planning database from 'P' to 'PL' to ensure that all records in the consolidated database had unique and searchable post-migration identifiers.
<p>Ensuring that you know the specific fields that will be migrated</p>	<ul style="list-style-type: none"> • Prior to migration, you or the contractors performing the migration will create an extract of data from your system to model how data will be removed, what it will look like post migration and where it will be presented in the new system. • It is very important to go through this extract thoroughly. It is likely that you will discover in the first iteration that not all data has been extracted accurately. For example, some picklists may not be carried through or some data elements may not have been picked up. Another common problem is for some data elements to be merged together, ie the extract may merge the

	title and notes fields. It is important that you identify all such problems and work out what is causing them prior to performing your migration.
--	---

System issues to consider in migration planning

<i>Issue:</i>	<i>Discussion:</i>
Metadata mapping	<p>Metadata mapping is a critical part of migration planning. Metadata mapping between the old and new systems will help to identify all your critical metadata and will make sure that it is all migrated. Ensuring all metadata can be migrated is necessary for maintaining functionality and can also be critical for ensuring that integrity and authenticity are preserved.</p> <ul style="list-style-type: none"> • Make sure you have a correct set of rules that define your metadata elements. It is very important to make sure that the meaning of elements is maintained across migrations. • If your organisation uses a range of identifiers for its information and if it needs to maintain a record of previous record identifiers, make sure these can be kept. • Make sure appropriate date detail is maintained. In some migrations the only date information carried across is the date of the migration. This becomes the default date for all recordkeeping actions, including date of creation. This severely impacts recordkeeping processes such as disposal and access and has significant impact on the useability and integrity of records. Therefore the meaning of each date field must be maintained following migration. • If you are using contractors to perform the migration, make sure that their metadata mapping is critical to signing off on your approval of the pre migration work they have done. • Business rules for the same function may be handled differently in the original and target systems. For example, access rights may be handled using caveats in one system and using user profiles in the other. It is necessary to plan a strategy for the migration of the required functionality using the different business rules.
Problems with current system functionality	<ul style="list-style-type: none"> • Remember that migration is an opportunity to improve system functionality. Look realistically at the problems and shortcomings in your current system. Don't simply bring them forward. Instead look at how you can use migration as an opportunity for system improvement.

Functionality of the target system	<ul style="list-style-type: none"> You need to ensure that you understand the full functionality of the target system. This includes recordkeeping functionality as well as other types of functionality.
Capacity issues in the target system	<ul style="list-style-type: none"> Systems can have different rules for the character length of fields. This can raise significant problems with migration. For example, in one organisation, their target system in its title field created a data string of the old creator and title fields. It also limited the title field to 100 characters. Following the migration, the first 100 characters of each record were taken up by the creator information (branch name and staff member name) and there was no space left for the actual record titles. This was completely unworkable and the migration had to be repeated which was costly, time consuming and led to further system downtime. Is there adequate server capacity for the target system? If the system is being rolled out to regional offices, is there sufficient bandwidth? Are response times adequate?
System interdependencies	<ul style="list-style-type: none"> If systems are talking to each other, what happens when one is upgraded? Are there live connections between the system requiring migration and the agency website or intranet? For example, is the latest version of a position description retained in the records system and linked live to the intranet? Do online transactional systems link to the business system being migrated? How can necessary functionality be supported post migration?
Issues associated with increased system functionality	<p>New target systems often have greater functionality than the systems they are replacing. In one organisation, data appeared in the new system that was not present in the old system. This was because the new system had greater capacity to be able to access data contained within the record. It is important to be aware of this possibility, identify the effects it will have and determine whether these are required by your organisation.</p>
Change management	<ul style="list-style-type: none"> You may need new procedure documentation if your new system is significantly different to its predecessor. User training is also critical so ensure that you have adequate time and resources for this.

Where relevant, plans should be made to address these or any other issues identified before your migration is conducted.

Tip: Consider the requirements of the *Standard on digital recordkeeping*

As official records, your migrated records should be managed in accordance with State Records' standards and requirements. Your migrated records need to be managed in systems that meet the requirements of the State Records *Standard on digital recordkeeping*.

Once your planning is complete you will have a thorough understanding of your records as well as your current and target systems. This will enable you to:

- begin to address the issues that must be resolved before migration can commence
- identify the desired target state of your records post migration, and
- develop or decide upon a migration method that will convert your records, including all metadata and essential characteristics, from their current state to the target state.

4.6 What about contractors?

Tip: What to do if contractors perform migrations on your behalf

In the contract you sign, you must specify exactly what migration tasks you want the contractor to perform on your behalf – this could include planning, identification of issues for remediation, metadata mapping, testing, implementation etc.

You must also ensure that you specify exactly what documentation you want from the contractor. For example you could request information concerning:

- all metadata mapping and documentation of issues requiring remediation
- all data cleansing and record consolidation
- all business rules applied in the original and target systems
- any user consultation
- any disposal performed and the authorisation for this action
- all plans for maintenance of connections between systems
- full migration testing (pre and post migration testing, including full quality assurances)
- full migration implementation, and
- appropriate sign off on the migration (remember ultimately responsibility for the outcomes of the migration process rests with your organisation).

You should also require full documentation of the capacities and functionality of the target system – the new system that your records have been migrated into. The structure of the target system and how it is implemented is particularly important – this will govern your day to day use of the system and will form the basis for your next system migration. System documentation you might request could include:

- system manuals
- table structures
- hardware configuration
- system procedures, etc.

4.7 Perform pre migration testing

Once you have developed your migration method and configured your target system, you need to perform a test migration on a small sample of duplicated records.

The resulting migrated records need to be assessed and verified by relevant technical and business staff to ensure that they are authentic reproductions, are complete, accessible and useable and that the migration strategy is appropriate. It is advisable that the team that validates both the pre and post migration testing is different to the team that actually designs and performs the migration.

Should adverse affects be noted in the migrated records, a revised migration strategy must be devised. This strategy should also be subject to pre migration testing.

Pre migration testing needs to be documented. This documentation should provide the basis for the final migration plan.

Once pre migration testing is complete, the pre migration testing and the finalised migration plan should be signed off by the Chief Information Officer or management official with appropriate authority.

4.8 Perform the migration

You are now ready to perform your migration.

Tip: Duplicate data prior to commencing migration

You may want to create a copy of the data you want to migrate just prior to migration. This gives you a full set of data to rely on if the migration results in the loss or corruption of data.

When performing your migration, you should aim to take records from their parent system and migrate them directly into their target system with as minimal intervention and time delay as possible. This helps to demonstrate an 'unbroken chain of custody' and can help prove the authenticity and reliability of the migrated records.

If you have the resources or if the records being migrated are significant corporate assets, you should undertake continuous quality assurance checking during the migration itself, followed by final verification of the success of the process at the end of the migration.

Performing the migration will also result in the creation of two copies of the same record – the original record or source record and the new migrated version of this record. When performing the migration, you need to ensure that the new migrated record is treated and used by staff as the new official record of business. The original or source record should not be accessible for staff to use. More guidance about the management of source records following their migration is provided below.

Remember, in a litigious environment or for high risk records, any real or perceived lapse in the management of the record during the migration process may be seen as compromising its authenticity.

Tip: Ensure system security

When performing migrations, you could consider the physical and logical security of the records. That is, you may want to control access to the physical space where the migration is being performed. You may also want to limit access to the platform in which the migration is being undertaken.

4.9 Perform post migration testing

Once your migration is complete, post migration testing must confirm that:

- all records requiring migration are migrated, including those that are stored off line, in non active systems or secondary storage environments
- the entire record, including all necessary metadata, is migrated
- all necessary business rules, functionality and essential characteristics have been preserved
- users are satisfied with the authenticity, completeness, accessibility and useability of the migrated record.

Tip: Involve relevant business staff

Involve the people who use the migrated records as part of their business operations to help you check that the migrations have been successfully performed. These people know the records well, know what is necessary to support their business operations and will be able to tell you whether the copies produced are adequate to meet their requirements.

This form of data checking should not be used as the sole measure of effective migration, but as a further means of validation.

Post migration testing does not have to be performed individually on every record – testing at an aggregate level is appropriate. The number of records selected for testing needs to be proportionate to the number of records maintained in the system.

If the migration proves to be inadequate, it must be repeated and the migration strategy redesigned and retested if required.

Example: Check and recheck the quality of your migration

A recent survey of newspaper archives in the United States highlights the importance of checking the quality of your migrated records and making sure that the records are backed up prior to migration and ensuring that the migration was performed adequately before source records are destroyed.

Of the organisations surveyed, only 18% had *not* experienced data loss during migration projects. (The losses identified ranged from minor – a few corrupt images on CD-ROMS – to disastrous – the complete loss of an entire collection of thousands of images.)⁴

It is very important to check your records that have been migrated so that the migration can be repeated if necessary. Validation of migration, through methods such as routine bit-level validation, is also necessary to help prove the ongoing authenticity of your records post migration. This is good data management practice and also potentially necessary if the records are to ever stand up as evidence in court.

All post migration testing needs to be documented.

Tip: Perform standard reports to verify the success of your migration

Develop a list of standard reports and standard searches that are common in your organisation. Do these searches in your original system and then your target system. You need to see the same results in both systems. The migration won't

⁴ This survey is discussed in *InterPARES 2 Project book, op.cit.*, p.289.

be successful until you can verify this. You should keep all these reports as part of your migration documentation as further validation of the work you have done. This documentation can also act as the basis for future migrations and will greatly reduce the planning preparations for these migrations.

Once post migration testing is complete, the migration process should be signed off by the Chief Information Officer or management official with appropriate authority.

4.10 Make records of your migration

The entire migration process and associated project planning should be documented. This could include:

- the records being migrated
- the trigger for the migration
- relevant research
- all decisions, including decisions not to migrate certain metadata components of a record
- risk assessments
- any disposal performed prior to migration
- the identified essential characteristics
- the technical requirements of the original and target systems
- the formal migration or transfer process
- the date and time of the migration and all personnel involved
- all system configurations, including metadata definitions and mappings
- all testing
- all reports that compare original system functionality to target system functionality
- all sign offs
- any data cleanup performed
- any variations to plans
- any necessary variation in records design, metadata, format or content that will or have resulted from the migration
- the disposal of the source records used, once the appropriate quality assurance period has elapsed.

4.11 Keep source records for at least six months

Following their successful migration source records, the records that were used as the input to the migration, **must be kept for at least six months**. Retaining the source records for at least this period will enable the migration to be repeated if it is discovered that some or all of the migrated records do not meet quality control standards or business requirements.

The six month retention period should begin to be calculated from the conclusion of successful post migration testing, where the migration and all outstanding issues associated with it have been signed off by the Chief Information Officer.

In high risk scenarios it is likely that source records will need to be kept for longer than the minimum six month retention period required by State Records' General

Retention and Disposal Authority – *Source records that have been migrated* (GA33).

If records need to be kept for more than six months public offices will need to use risk assessments to determine an appropriate retention period. These risk assessments should at a minimum consider:

- the business purpose of the records
- the risks associated with this business
- the potential business, financial and legal implications of the loss of or damage to the migrated records
- the size and complexity of the migration and the likelihood of problems associated with it
- the complexity of the records being migrated
- the capacities of the target system and the possibility that all aspects of this system and its impact on the records are not fully understood at the time of migration
- the state of the original system and its compliance with standards and best practice requirements, including metadata standards and other business requirements.

Public offices must always err on the side of caution and if any problems or concerns with the migrated or transferred records are noted, such as corruption of portions of the record, or loss of information or distortions in the records caused by the new capacities of the target system, then the identified retention period must be extended by at least another six months. Again, this additional retention period should begin to be calculated from the conclusion of successful post migration testing, where the repeated migration and all outstanding issues associated with it have been signed off by the Chief Information Officer.

The retention of source records for the short or long term needs to be planned and effectively managed. During their retention period following migration, source records should be stored and protected to ensure that they remain as accountable, well managed records of business that can be used again as appropriate source records should the entire migration or portions of it need to be repeated. However, during their retention period, it must be remembered that these records are no longer the official records of business. They must be protected against user access and must not be used in business transactions.

It is a risk based business decision for the public office to determine whether the source records should be stored as flat files in the target system or actively maintained within their original system with no user access for the identified retention period prior to their destruction.

If it is determined that the source records can be destroyed, this must be documented and signed off by the delegated officer, usually the Chief Information Officer.

Further assistance:

For further advice about the issues associated with the retention of source records see the General Retention and Disposal Authority - *Source records that have been migrated* (GA33) at [http://www.records.nsw.gov.au/recordkeeping/source_records_\(ga33\)_15630.asp](http://www.records.nsw.gov.au/recordkeeping/source_records_(ga33)_15630.asp).

4.12 Think strategically about migration

Don't just see migration as a necessary records management and ICT process. Instead see it as a business enhancement and business improvement project. Migration is about improving accessibility and useability and is therefore directly related to the organisational bottom line. See it and market it as this type of project.

In terms of improving recordkeeping practice, migration can be an opportunity to:

- build better automation (for example, automated metadata capture, sentencing at creation etc)
- better integrate recordkeeping tools into the business system
- implement more appropriate security controls
- implement automated workflow
- include additional attributes to improve searching and user access
- improve data sharing and interfaces between systems, or take advantage of the upgraded functionality provided by the system you are migrating to.

4.13 Checklist for migration

Plan for migration

<i>Has your organisation...</i>	Yes	No
Identified migration triggers that signify a migration needs to be performed? Common triggers include: <ul style="list-style-type: none"> • the adoption of new business systems • impending technological change or obsolescence • the need to move high value records to stable long term formats • administrative change • the adoption of shared service arrangements • outsourcing of business functions • the need to transfer records to State Records. 		
Understood the unique characteristics of records that need to be preserved during migration? These are: <ul style="list-style-type: none"> • record complexity • metadata • essential characteristics. 		
Determined the type of migration required? This could be: <ul style="list-style-type: none"> • minor functionality change • medium levels of functionality change • migration to a complete new system. 		
Understood the different levels of planning and implementation required by the type of migration that is planned?		

<p>Appropriately planned for all relevant data issues in the migration? These could include:</p> <ul style="list-style-type: none"> • performing data cleaning • implementing a records clean up • maintaining all appropriate metadata • destroying records where appropriate • identifying independent systems that need to be captured into an organisational recordkeeping system and migrated • understanding the hardware, software and format requirements of the records requiring migration • identifying non standard records • checking older records that have undergone multiple migrations • identifying all records requiring migration, including records in online and offline storage, records in all business environments, including shared service environments, and secondary storage environments • initiating data cleansing • performing necessary record consolidation • planning for the management of aspects of record context that are difficult to manage through migration, such as movement histories • making implicit metadata explicit • talking to users to ensure that all their business requirements for records can be met by the proposed migration • managing records that are encrypted or contain digital signatures • checking the quality of encoding schemes • verifying user understanding of existing metadata fields • determining whether multiple values have been used in single metadata fields • ensuring existing count downs remain in place • comprehensively assessing metadata use in the system • verifying use of date information • identifying all records across databases that will be consolidated • verifying which fields will be migrated and how and where. 		
<p>Appropriately planned for all relevant system issues in the migration? These could include:</p> <ul style="list-style-type: none"> • completing an extensive metadata mapping, to ensure that all necessary metadata elements, their corresponding meaning and functionality, relevant business rules and links to other systems can be migrated between systems 		

<ul style="list-style-type: none"> • assessing how current functionality could be improved • ensuring you understand the capacities of the target system • assessing capacity issues in the target system • identifying interdependencies between systems that need to be maintained • identifying potential issues with the increased functionality of the target system. 		
Addressed the issues that need to be resolved before migration can commence?		
Configured the target system so that it meets all your business requirements?		
Identified the desired target state of the records post migration?		
Developed a migration method that will convert the records, including all metadata and essential characteristics, from their current state to the target state?		
<p>Identified what you will require from any contractors who perform migration on your behalf? This could include all information concerning:</p> <ul style="list-style-type: none"> • all metadata mapping and documentation of issues requiring remediation • all data cleansing and record consolidation • all business rules applied in the original and target systems • any user consultation • any disposal performed • all plans for maintenance of connections between systems • full migration testing, and • full migration implementation. 		

Undertake pre migration testing

<i>Has your organisation...</i>	Yes	No
Performed a test migration on a small sample of duplicated records?		
Had test migrated records assessed by relevant technical and business staff to determine their adequacy?		
Redeveloped its migration strategy and performed test migrations again if adverse affects were noted in the initial pre migration testing?		
Documented all pre migration testing?		
Finalised your migration plan based on the results of your pre migration testing?		

Had the pre migration testing and finalised migration plan signed off by the CIO when the pre migration testing yielded complete, accessible and useable copies of the records?		
---	--	--

Perform the migration

<i>Has your organisation...</i>	<i>Yes</i>	<i>No</i>
Performed the complete migration?		

Perform post migration testing

<i>Has your organisation...</i>	<i>Yes</i>	<i>No</i>
Ensured that all records have been migrated?		
On a sample that represents an adequate proportion of the total number of records migrated, ensured and tested that: <ul style="list-style-type: none"> complete records, including all necessary metadata, has been migrated? all necessary business rules have been retained? all necessary functionality has been retained? all essential characteristics have been retained? users are satisfied with the authenticity, completeness, accessibility and useability of the migrated records? 		
Repeated the migration, if testing reveals that any problems with the migration have occurred? If problems are noted, the migration process will need to be redesigned and pre migration testing performed again.		
Documented all post migration testing?		
Had the completed migration signed off by the CIO or management official with appropriate authority, when the post migration testing yields complete, accessible and useable copies of the records?		

Make records of migration

<i>Has your organisation...</i>	<i>Yes</i>	<i>No</i>
Documented migration planning, testing and implementation? This could include: <ul style="list-style-type: none"> the records being migrated the trigger for the migration relevant research all decisions risk assessments the identified essential characteristics the technical requirements of the original and target systems 		

<ul style="list-style-type: none"> • the formal migration or transfer process • the date and time of the migration and all personnel involved • all system configurations, including metadata definitions and mappings • all testing • all reports that compare original system functionality to target system functionality • all sign offs • any data cleanup performed • any variations to plans • any necessary variation in records design, metadata, format or content that will result from the migration • the disposal of the source records used. 		
---	--	--

Ensure source records are kept for appropriate periods of time

<i>Has your organisation...</i>	<i>Yes</i>	<i>No</i>
Performed risk assessments to determine how long source records should be retained?		
Stored source records in an appropriate environment where they: <ul style="list-style-type: none"> • cannot be accessed by users? • cannot be mistaken for the official version of the records? • are secure? • can be accessed by technical staff in order to repeat all or portions of the migration as required? 		
Repeated the migration using the source records if it is found that some or all of the migrated records do not meet quality control standards or business requirements?		
Documented the destruction of the source records, if they have been destroyed?		

© State of New South Wales through the State Records Authority, February 2009. This work may be freely reproduced and distributed for most purposes, however some restrictions apply. See the copyright notice on www.records.nsw.gov.au or contact State Records.
 ISBN: 978-0-9805148-9-6