Data Migration

WHAT RECORDS MANAGERS NEED TO KNOW

March 2019

Nitza Medina-Garcia, CRM
Introduction

- More and more records and information are stored electronically
- Records and information need to be moved periodically from one software program to another
- Records Managers need to understand the typical software project phases
- Records Managers need to understand the data migration process
What is Data Migration?

Data migration is the process of transferring data between storage types, formats, or computer systems. It is a key consideration for any system implementation, upgrade or consolidation.

From Wikipedia
Drivers of Data Migration

- Mergers and acquisitions
- Library consolidation efforts
- Large number of people coming or going
- Company / department dispositions
- Additional software functionality needs
- Information governance initiatives (fileshare and email cleanups)
Benefits of Data Migration

- Remove duplication of effort
- Utilize additional software features
- Remove duplication of content
- Remove silos of information
Software Project Phases (Waterfall)

- Project Kickoff
- Discovery and Requirements
- Software Design
- Software Build
- Data Migration
- User Acceptance Testing
- Training
- Go Live
Software Project Phases (Agile)

- Project Kickoff
- Discovery and Requirements
- Data Migration
- Software Build
- User Acceptance Testing
- Iteration 1, 2, 3 and so on
- Approval
- Training
- Go Live
Data Migration Tasks

• Data Migration Strategy
• Identification of Data Sources
• Analysis of Data Sources
• Data Clean Up
• Data Conversion
• Data Migration Execution
• Data Validation
Records Appraisal Tasks

• Appraisal Strategy
• Identification of Record Sources
• Analysis of Records
• Records Clean Up
• Records Conversion
• Records Move
• Records Approval
Key Reasons for Data Migration Failure

- Technocentricty
- Lack of specialist skills
- Underestimation
- The blame game
Golden Rules for Data Migration

1. Data Migration is a business not a technical issue
2. The business knows best
3. No organization needs, wants or will pay for perfect quality data
4. If you can’t count it, it doesn’t count
Best Practices for Data Migration

1. Create a Migration Strategy
2. Data Inventory and Analysis
3. Gap Analysis and Mapping
4. Data Migration Execution
5. Data Validation
6. Legacy Software Decommissioning
1. Create a Migration Strategy

- Determine what data sources are in/out of scope
- Determine if following an agile or waterfall methodology
- Determine key stakeholders and team members
- Communicate migration strategy
2. Data Inventory and Analysis

Is there one data source or multiple?
How big are the data sources?
Where are the data sources located?
Is there ROT (redundant, obsolete, trivial) data that needs to be deleted?
Will there be a standardization on a certain set of codes?
Who will have the authority to make data decisions?
3. Gap Analysis and Mapping

<table>
<thead>
<tr>
<th>RMS Table.Field Name</th>
<th>Field Type [# characters]</th>
<th>RMS Caption</th>
<th>Old DB Table.Field Name</th>
<th>Old DB Caption Name</th>
<th>Rules</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>File.id</td>
<td>Varchar[15]</td>
<td>ID</td>
<td>Jacket.id</td>
<td>ID</td>
<td>Make sure there are 5 digits, if not add leading zeros</td>
<td></td>
</tr>
<tr>
<td>File.name</td>
<td>Varchar[122]</td>
<td>Name</td>
<td>Jacket.name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>File.desc</td>
<td>Varchar[2048]</td>
<td>Description</td>
<td>Jacket.desc</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>File.resp_user</td>
<td>Trustee</td>
<td>Responsible User</td>
<td>Jacket.user1</td>
<td>Responsible User</td>
<td>User list will come from AD; legacy users will come from Old DB</td>
<td></td>
</tr>
<tr>
<td>File.office</td>
<td>Metadata Table</td>
<td>Office</td>
<td>Jacket.office</td>
<td>Office</td>
<td>***Need list of Office codes</td>
<td></td>
</tr>
<tr>
<td>File.dept</td>
<td>Metadata Table</td>
<td>Department</td>
<td>Jacket.dept</td>
<td>Department</td>
<td></td>
<td></td>
</tr>
<tr>
<td>File.open_date</td>
<td>Date Field</td>
<td>Open Date</td>
<td>Jacket.odate</td>
<td>Open Date</td>
<td>Convert to UTC</td>
<td></td>
</tr>
<tr>
<td>File.close_date</td>
<td>Date Field</td>
<td>Close Date</td>
<td>Jacket.cdate</td>
<td>Close Date</td>
<td>Convert to UTC</td>
<td></td>
</tr>
<tr>
<td>File.comments</td>
<td>Varchar[4000]</td>
<td>Comments</td>
<td>Jacket.notes[1-5]</td>
<td>Comments</td>
<td>Concatenate lines 1-5 into new Comments field. Any comments over 50 characters will be truncated.</td>
<td>***Verify the largest amount of data in Old DB</td>
</tr>
</tbody>
</table>
4. Data Migration Execution

- Agreed rules used to create data migration scripts
- Data extracted from data source (old DB)
- Data stored in staging table
- Data imported into new software database
## 5. Data Validation

<table>
<thead>
<tr>
<th>#</th>
<th>Issue Name</th>
<th>Description</th>
<th>Reporter</th>
<th>Steps to Recreate</th>
<th>Priority</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Missing 10 folders</td>
<td>Folders 1-10 do not appear in new db</td>
<td>Julia Rivera</td>
<td>1. Search for folders 1-10 individually</td>
<td>High</td>
<td>Check extraction database to see if information is missing; check to see if folders were created by user after extraction date May 6, 2016.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. See no results appear</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Folder Comments Missing</td>
<td>Folder comments only include 10 characters. The rest of the comments are missing.</td>
<td>JohnGills</td>
<td>1. Search for Folder 32</td>
<td>High</td>
<td>Check extraction scripts and database.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2. Open folder properties.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3. See comments are partially missing</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5. Legacy Software Decommissioning

• Determine **when** the old database software should be turned off

• Two strategies:
  • Double data entry (old and new system) for a short period
  • Delta migration over a short period (weekend)
Key Take Aways

Data migration is not solely the responsibility of IT.
Records Managers must become familiar with the data migration process.
Allocate enough time/buffer in data migration project.
Set right expectations for all involved in the beginning.
Good Resources

DAMA International (organization)
www.dama.org

Data Migration Pro (community)
www.datamigrationpro.com
Questions?
InfoCompass Business Solutions

Web: https://infocompass.com
Blog: https://infocompass.com/blog/
Email: nmg@icbsolutions.net

Nitza Medina-Garcia, CRM

Please contact me if you have any further questions