

Data Migration Roadmap Guidance

Version 3.2 • 7/9/2019

Final

Document Version Control

VERSION	DATE	AUTHOR	DESCRIPTION
1.0	5/3/07		Final changes
1.1	9/16/08		Revisions (removed Best Practices)
2.0	2/17/12		Reformats and modifications across the board
2.1	3/9/2012		Changes resulting from review
2.2	4/3/13		Updated the FSA Logo and inserted quick parts for Document Version, Document ID, Version Date, and updated headers/footers during audit.
3.0	2/3/2017	EITA-SME	Converted to new FSA Master Template format and included include changes from a EDM Standards quality assessment.
3.1	2/27/2017	EITA-SME	Repaired document paging issue, updated Appendices, updated title to include "Guidance" and improved TOC.
3.2	7/9/2019	Mike Fillinich	Inserted section 3.4.9.2 and added same language to line 4.1 of the Data Migration Project Deliverables Checklist.

Table of Contents

SECTION 1. INTRODUCTION	1
1.1. BACKGROUND AND PURPOSE	1
1.2. BENEFITS OF A DATA MIGRATION ROADMAP	1
1.3. INTENDED AUDIENCE	1
1.4. REFERENCE DOCUMENTS / APPLICABLE PROJECT DOCUMENTS	1
SECTION 2. EXECUTIVE SUMMARY	3
SECTION 3. DATA MIGRATION ROADMAP	5
3.1. INTRODUCTION	5
3.1.1. <i>Data Migration Project Lifecycle</i>	6
3.2. DATA MIGRATION PLANNING PHASE	7
3.2.1. <i>Planning Overview</i>	7
3.2.2. <i>Data Migration Planning Tasks & Subtasks</i>	8
3.2.3. <i>Planning Data Migration Project</i>	9
3.2.3.1. Establish Scope	9
3.2.3.2. Identify Risk / Constraints / Dependencies / Assumptions	9
3.2.3.3. Develop Data Migration Risk Mitigation Plan	10
3.2.3.4. Develop Data Migration Communications Plan.....	12
3.2.3.5. Critical Success Factors for Data Migration Planning	13
3.2.4. <i>Determine Data Migration Requirements</i>	14
3.2.4.1. Determine Business Requirements and Expectations	14
3.2.4.2. Determine Technology and IT Infrastructure Requirements	15
3.2.4.3. Identify Relevant Best Practices	15
3.2.5. <i>Assess Current Environment</i>	16
3.2.5.1. Identify and Collect Existing Data related Architecture	16
3.2.5.2. Blueprint Current Stat of the Data Architecture	17
3.2.5.3. Determine Data Migration Technology.....	17
3.2.6. <i>Develop Data Migration Plan</i>	17
3.2.6.1. Determine Data Migration Method	17
3.2.6.2. Determine Data Conversion Plan.....	18
3.2.6.3. Determine Data Integration Plan.....	18
3.2.6.4. Plan Parallel Operation	19
3.2.6.5. Develop Migration Data Quality Plan	19
3.2.6.6. Develop Data Archival Strategy	20
3.2.6.7. Develop Data Migration Test Plan	20
3.2.7. <i>Define and Assign Roles & Responsibilities</i>	20
3.2.7.1. Define Migration Roles and Responsibilities	20
3.2.8. <i>Data Migration Planning Deliverables</i>	21
3.2.9. <i>Data Migration Planning Checklist</i>	22
3.3. DATA MIGRATION ANALYSIS AND DESIGN.....	22
3.3.1. <i>Analysis and Design Overview</i>	22
3.3.2. <i>Data Migration Analysis and Design Tasks and Subtasks</i>	22
3.3.3. <i>Perform Data Migration Analysis</i>	23
3.3.3.1. Analyze Current Environment	23
3.3.3.2. Evaluate Data Migration Technology	23
3.3.3.3. Evaluate Data Quality	24
3.3.3.4. Perform Data Profiling.....	24
3.3.3.5. Critical Success Factors for Data Migration Analysis & Design	25
3.3.4. <i>Determine Data Security Controls</i>	26
3.3.4.1. Determine Enterprise Management and Operational Security Controls.....	26
3.3.5. <i>Design Data Migration Environments</i>	26
3.3.5.1. Design Staging Area	26
3.3.5.2. Design Target Data Architecture	27
3.3.5.3. Correlate Migration Data (Source/Staging/Target).....	27
3.3.6. <i>Design Data Migration Procedures</i>	27
3.3.6.1. Design Data Staging Procedures	27
3.3.6.2. Design Data Cleansing Procedures	27
3.3.6.3. Design Data Conversion Procedures	28

3.3.6.4. Design Target Data Migration Procedures	28
3.3.6.5. Design Data Validation Procedures	28
3.3.6.6. Design Data Quality Remediation Procedures	28
3.3.6.7. Refine Data Migration Test Plan	28
3.3.7. <i>Data Migration Analysis & Design Deliverables</i>	28
3.3.8. <i>Design Checklist</i>	29
3.4. IMPLEMENTATION	30
3.4.1. <i>Implementation Overview</i>	30
3.4.2. <i>Data Migration Implementation Tasks & Subtasks</i>	30
3.4.3. <i>Develop Data Migration Procedures</i>	31
3.4.3.1. Configure Resources	31
3.4.3.2. Develop and Test Data Migration Procedures	32
3.4.3.3. Develop and Test Data Validation Procedures	32
3.4.3.4. Develop and Test Data Cleansing Procedures	32
3.4.3.5. Develop and Test Data Conversion/Transformation Procedures	32
3.4.3.6. Establish Access to Staging and Target Areas	32
3.4.3.7. Critical Success Factors for Implementation	32
3.4.4. <i>Stage Data</i>	33
3.4.4.1. Create Staging Area	33
3.4.4.2. Populate Staging Area	33
3.4.4.3. Integrate Staged Data	33
3.4.4.4. Validate Staged Data	33
3.4.5. <i>Cleanse Data</i>	33
3.4.5.1. Cleanse Data According to Data Remediation Plan	33
3.4.5.2. Validate Cleansed Data	34
3.4.6. <i>Convert / Transform Data</i>	34
3.4.6.1. Convert / Transform Data	34
3.4.6.2. Validate Converted / Transformed Data	34
3.4.7. <i>Migrate Data</i>	34
3.4.7.1. Perform Trial Data Migration	34
3.4.7.2. Validate Results of Trial Migration	34
3.4.7.3. Obtain Approval for Full Migration into Production Environment	34
3.4.7.4. Perform Full Data Migration--Deployment	34
3.4.7.5. Validate Results of Full Migration	35
3.4.8. <i>Post-Migration</i>	35
3.4.8.1. Operate Legacy & Target Environments in Parallel	35
3.4.8.2. Validate Parallel Operation	35
3.4.8.3. Release Data Environments	35
3.4.9. <i>Data Migration Implementation Artifacts</i>	35
3.4.9.1. Implementation Checklist	36
3.5. DATA MIGRATION CLOSEOUT	36
3.5.1. <i>Closeout Overview</i>	36
3.5.2. <i>Data Migration Closeout Tasks & Subtasks</i>	36
3.5.3. <i>Document Data Migration Results</i>	36
3.5.3.1. Critical Success Factors for Data Migration Closeout	36
3.5.4. <i>Document Data Migration Lessons Learned</i>	37
3.5.5. <i>Perform Knowledge Transfer</i>	37
3.5.6. <i>Perform Stakeholder Training</i>	37
3.5.7. <i>Disseminate Data Migration Results</i>	37
3.5.8. <i>Obtain Approval for Project Completion</i>	37
3.5.9. <i>Data Migration Closeout Artifacts</i>	37
3.5.10. <i>Data Migration Closeout Checklist</i>	37
APPENDIX A - ACRONYMS AND ABBREVIATIONS	A-1
APPENDIX B - GLOSSARY	B-1
APPENDIX C - DATA MIGRATION PROJECT REVIEW CHECKLIST	C-1
APPENDIX D - DATA MIGRATION PROJECT DELIVERABLES	D-1

List of Figures

Figure 2-1: Data Migration Project Lifecycle	3
--	---

Figure 3-1: High-Level Data Migration Process.....	5
Figure 3-2: Scope of the Data Migration Project	5
Figure 3-3: Data Migration Project Lifecycle	6
Figure 3-4: Data Management Activities in Data Migration	7
Figure 3-5: Data Migration Implementation Plan Flow	30

List of Tables

Table 1-1: Reference Documents	2
Table 3-1: Data Migration Lifecycle with High-Level Tasks Identified	7
Table 3-2: Data Migration Planning	8
Table 3-3: Assumptions, Constraints, and Risks.....	10
Table 3-4: Data Migration Risk Probability and Impact Levels	11
Table 3-5: Data Migration Risk Mitigation Matrix.....	12
Table 3-6: Data Migration Roles and Responsibilities.....	21
Table 3-7: Data Migration Analysis and Design	23
Table 3-8: Data Migration Implementation	31
Table 3-9: Data Migration Closeout.....	36
Table A-1: Acronyms and Abbreviations	A-1
Table B-1: Glossary	B-2
Table C-1: Data Migration Review Checklist.....	C-4
Table D-1: Data Migration Deliverables	D-4

Section 1. Introduction

1.1. Background and Purpose

Federal Student Aid is engaged in a long-term effort to integrate its processes, data and systems. To better support these business objectives and to emphasize data as an enterprise asset, Federal Student Aid has established the Enterprise Data Services (EDS) team. The goal of the EDS is to consistently define data and make standardized data available across the enterprise by providing information services and data technology expertise to business owners, project managers and architects.

This document outlines a roadmap and provides checklists to assist with mitigating some of these challenges. Comments or suggestions for improvement to this roadmap are encouraged and should be reported back to the Project Manager for Enterprise Data Management.

1.2. Benefits of a Data Migration Roadmap

One way to increase the chances of success on a data migration project is to establish and follow a framework, which is based on best practices, tested and improved through lessons learned. Use of this roadmap, in addition to a project lead's experience, is expected to increase the quality of data migration projects. Specifically, the following benefits are expected:

- **Minimal disruption to the business:** The key to minimal disruption is thorough planning and coordination. Planning should include a timeline for each stage of the migration. In addition, coordination with business stakeholders throughout each stage is important.
- **Efficient resource utilization (people, budget and time):** Proper data migration planning allows for efficient use of resources. In addition, this planning sets up performance metrics, which can be measured and used to make decisions throughout the migration.
- **Quality Assurance and Risk Mitigation:** Establishing and executing quality assurance and risk mitigation throughout the data migration project improves the chances of a successful project. Proactive quality assurance manages the quality of artifacts and outputs produced and early risk mitigation reduces negative impacts to the project.
- **Cost reduction:** Following a best practice roadmap is expected to result in an overall cost reduction. Also, minimal business disruptions should result in reduced cost.

1.3. Intended Audience

The target audience is individuals not familiar with data migration, and this document can serve as a reference for those familiar with the topic.

1.4. Reference Documents / Applicable Project Documents

The following external documents provide either governance or guidance for this document.

DOCUMENT ID	DOCUMENT TITLE	DOCUMENT VERSION
	Front page design and single-line logo drawn from the <i>Department of Education Federal Student Aid Style Guide</i> .	
	Outline format drawn from the format of the U.S. Department of Education Enterprise Data Architecture – Enterprise Data Standards and Roadmaps.	
	Creative Computing Solutions, Inc. Enterprise Data Management (Operations) Technical Proposal	August 29, 2006

DOCUMENT ID	DOCUMENT TITLE	DOCUMENT VERSION
	Enterprise Data Management (Operations) Statement of Work	August 25, 2006
	Burry, Christopher & Mancusi, David. "How to plan for data migration." ComputerWorld	May 21, 2004
	Softek, Inc. 2006 Best Practices for Data Migration (White Paper). KnowledgeStorm.com: Softek, Inc., 2006.	
	Softek, Inc. Simplifying Technology Refresh with Data Migration Software (White Paper). KnowledgeStorm.com: Softek, Inc	January 2006
	Softek, Inc. The Hidden Costs of Data Migration (White Paper). KnowledgeStorm.com: Softek, Inc.	April 1, 2006
	Lewis, William. Resource Conversion (MS PowerPoint Presentation). Boston, MA: EMELD	January 9, 2004
	Peipert, Glenn & Cohen, Lori. Strategic Approach to Data Migration (White Paper). WWW: Conversion Services, International, Inc.	July 2005
	Manek, Parul. Microsoft CRM Data Migration Framework (White Paper). WWW: Microsoft, Inc	April 2003
	Damolakis, James. "Best Practices." Storage Magazine	November 2006
	Wikipedia- Best Practice	
	Wikipedia- Data Conversion	
	Wikipedia- Data Migration	
	Purba, Sanjiv. Handbook of Data Management 1999. Washington, D.C.: Auerbach	1999
	Inmon, W.H. Data Architecture: The Information Paradigm. Boston: QED Technical Publishing Group	1992
	Parthasarathi, Arvind. "Taking the Pain Out of Data Migration: Methodology that Works". DMDirect Newsletter	November 4, 2005
	Perot Systems, Integrated Partner Management (IPM) – Data Management Plan	Spring 2007

Table 1-1: Reference Documents

Section 2. Executive Summary

This document is the result of best practice research regarding data migration. It explains what data migration is, the steps involved, common problems and possible risks. This document outlines a practical roadmap to assist with the management of data migration projects.

The Enterprise Data Services (EDS) Team commissioned this document as a service for business representatives involved in data migration projects. It provides a high-level overview for those individuals not familiar with data migration and can also serve as a reference for those familiar with the topic. This document is written in business language and while there are technical components, it is designed for Project Managers, Subject Matter Experts, and non-technical staff.

Data migration is the transfer of data from one location, storage medium, or hardware/software system to another. Migration efforts are often prompted by the need for upgrades in technical infrastructure or changes in agency business requirements.

A review of best practices found two principles inherent in successful data migration efforts:

1. Perform data migration as a project dedicated to the unique objective of establishing a new (target) data store.
2. Perform data migration in four primary phases: **Data Migration Planning**, **Data Migration Analysis and Design**, and **Data Migration Implementation**, and **Data Migration Closeout** as shown in Figure ES-1.

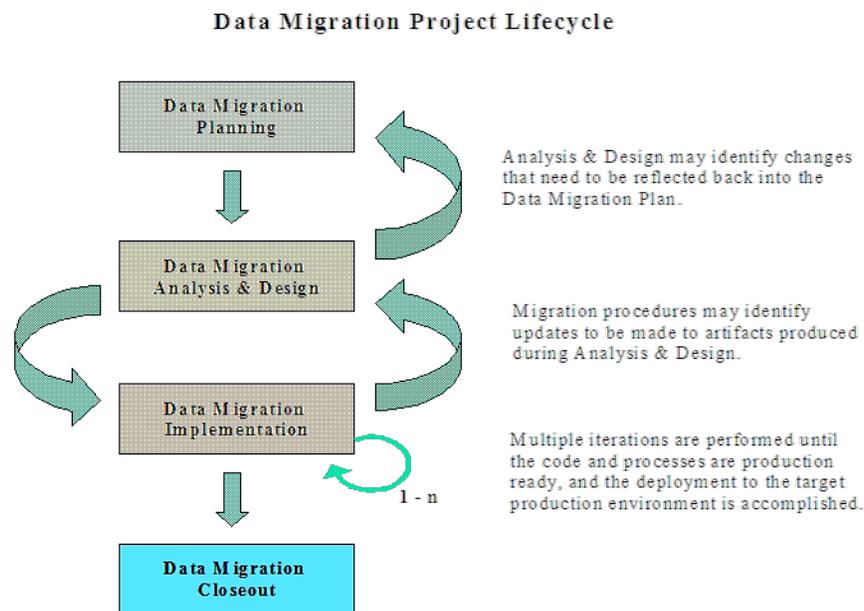


Figure 2-1: Data Migration Project Lifecycle

In addition, research found that the most successful projects were ones that maximized opportunities and mitigated risks. The following critical success factors were identified:

- Perform data migration as an independent project¹
- Establish and manage expectations throughout the process
- Understand current and future data and business requirements
- Identify individuals with expertise regarding legacy data.²

¹ Microsoft CRM Data Migration Framework, page 6

² Microsoft CRM Data Migration Framework, page 8

- Collect available documentation regarding legacy system(s).
- Define data migration project roles & responsibilities³ clearly.
- Perform a comprehensive overview of data content, quality, and structure.⁴
- Coordinate with business owners and stakeholders to determine importance of business data and data quality.

This document is organized according to the four primary phases: **Data Migration Planning**, **Data Migration Analysis and Design**, **Data Migration Implementation**, and **Data Migration Closeout** and contains a detailed description of each phase (including tasks and subtasks). In addition, common pitfalls are identified and described. Finally, this document contains a Data Migration Review Checklist, which serves as a tool to help launch and manage data migration projects

³ Microsoft CRM Data Migration Framework, page 7

⁴ *Strategic Approach to Data Migration*, page 3

Section 3. Data Migration Roadmap

3.1. Introduction

A data migration project focuses on the movement of data between legacy (source) data system(s) and a target system, including all necessary procedures for transferring and validating the data throughout the entire process (see Figure 2-1). Before data is moved, often it needs to be modified and/or transformed. This process is called Data Conversion. Planning and performing data conversion requires the development of transformation rules and procedures to implement the necessary changes. For example, if the legacy system stores date information in text format but the target system requires this information to be stored as date format, then a conversion of the legacy data is necessary prior to the data migration.

It is common to use a staging area as an interim data store to facilitate testing and validation of these modifications/transformations. In addition, a staging area can serve as a storage area for integration projects, which pull data from multiple source systems.

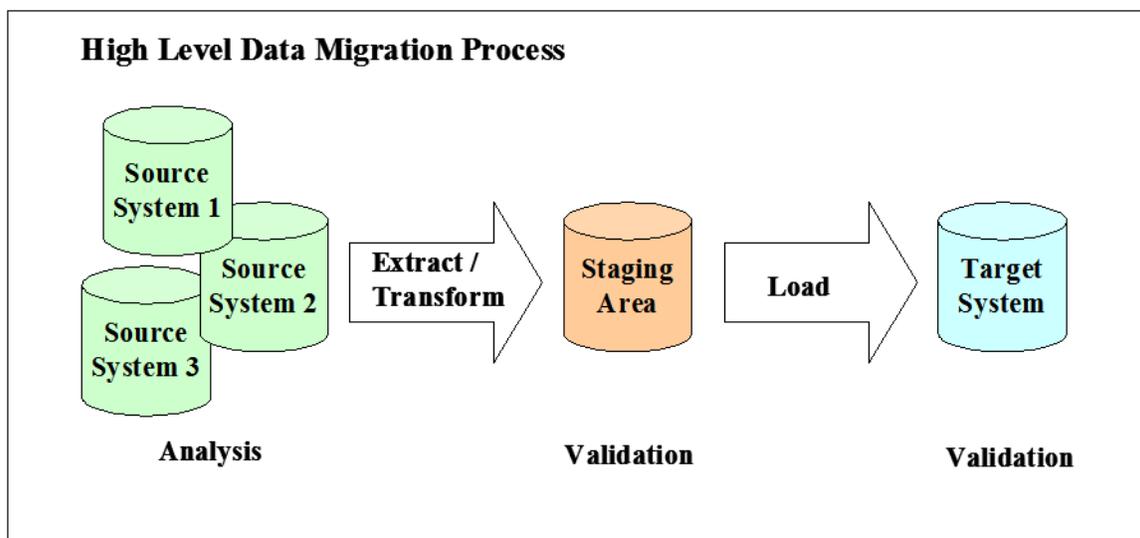


Figure 3-1: High-Level Data Migration Process

A review of best practices produced the following two principles inherent in successful data migration:

Perform data migration as a project dedicated to the unique objective of establishing a new (target) data store.

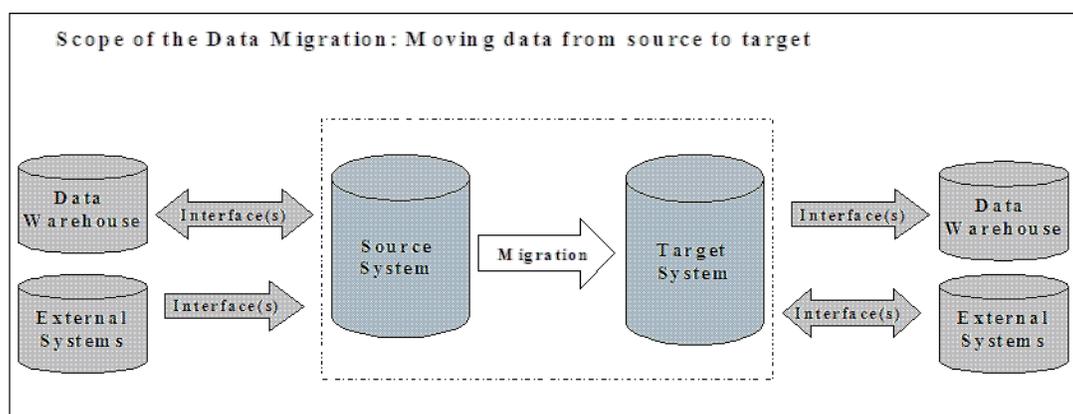


Figure 3-2: Scope of the Data Migration Project

Perform data migration in four primary phases as shown in Figure 2-3:

- Data Migration Planning
- Data Migration Analysis and Design
- Data Migration Implementation
- Data Migration Closeout

Data Migration Project Lifecycle

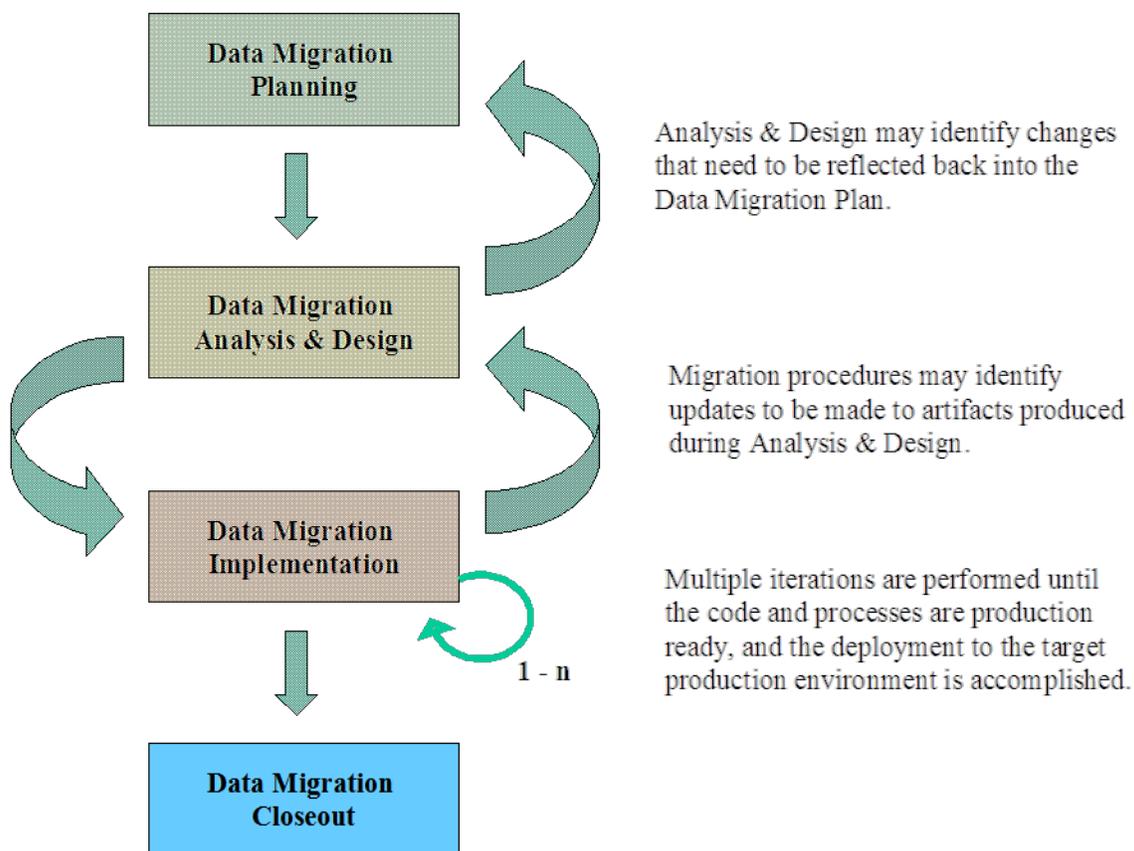


Figure 3-3: Data Migration Project Lifecycle

3.1.1. Data Migration Project Lifecycle

Table 2-1 lists the high-level processes recommended for each phase of the Data Migration Project Lifecycle. While all data migration projects follow the four phases in the Data Migration Project Lifecycle, the high-level and low-level processes may vary depending on the size, scope and complexity of each migration project. Therefore, the following information should serve as a guideline for developing, evaluating, and implementing data migration efforts. Each high-level and low-level process should be included in a Data Migration Plan.

DATA MIGRATION PLANNING PHASE	DATA MIGRATION ANALYSIS & DESIGN PHASE	DATA MIGRATION IMPLEMENTATION PHASE	DATA MIGRATION CLOSEOUT PHASE
Plan Data Migration Project	Analyze Assessment Results	Develop Procedures	Document Data Migration Results
Determine Data Migration	Define Security Controls	Stage Data	Document Lessons

DATA MIGRATION PLANNING PHASE	DATA MIGRATION ANALYSIS & DESIGN PHASE	DATA MIGRATION IMPLEMENTATION PHASE	DATA MIGRATION CLOSEOUT PHASE
Requirements			Learned
Assess Current Environment	Design Data Environment	Cleanse Data	Perform Knowledge Transfer
Develop Data Migration Plan	Design Migration Procedures	Convert Transform Data (as needed)	Communicate Data Migration Results
Define and Assign Team Roles and Responsibilities	Validate Data Quality	Migrate Data (trial/deployment)	
		Validate Migration Results (iterative)	
		Validate Post-Migration Results	

Table 3-1: Data Migration Lifecycle with High-Level Tasks Identified

During the lifecycle of a data migration project, the team moves the data through the activities shown in Figure 2-4.



Figure 3-4: Data Management Activities in Data Migration

The team will repeat these data management activities as needed to ensure a successful data load to the new target data store.

3.2. Data Migration Planning Phase

3.2.1. Planning Overview

The Data Migration Planning Phase describes the individual tasks to:

- Plan Data Migration Project
- Determine Data Migration Requirements
- Assess Current Environment
- Develop Data Migration Plan
- Define and Assign Team Roles and Responsibilities

To ensure that both the data migration project and the larger development project are successful, it is good practice to execute data migration as an independent project. Thorough planning is the foundation for consistent success in any process, and data migration is no exception. Also, a successful data migration effort requires the mitigation of issues and risks to the business/ organization.

The *Data Migration Plan* details the information that should be included for each step in the plan. Other results, such as risks and/or critical success factors, may simply be documented in the plan. All steps within subsequent phases of the migration are included in the plan. Also included are the way in which the steps should be performed with respect to rules, parameters, and procedures, and so forth.

In addition, the development program describes the general project deliverables, to which all projects must adhere, such as the Quality Plan, Change Management Plan, and Communications Plan.

3.2.2. Data Migration Planning Tasks & Subtasks

The following table presents the five major tasks of the Data Migration Planning Phase. Each task is broken down further into subtasks. Subtasks may occur in parallel, or in the sequence shown in this chart. The Data Migration Project Manager, in collaboration with the Data Migration Team (DMT), will determine the order in which these tasks will be performed.

DATA MIGRATION PLANNING PHASE				
Plan Data Migration Project	Determine Data Migration Requirements	Assess Current Environment	Develop Migration Plan	Define and Assign Roles & Responsibilities
Establish Scope	Determine Business Requirements & Expectations	Identify and Collect Existing Data-Related Artifacts	Determine Data Migration Method	Define Migration Roles and Responsibilities
Identify Risks/Constraints/Dependencies /Assumptions	Determine Technology and IT Infrastructure Requirements	Blueprint Current State of the Data Architecture	Determine Conversion Plan	
Develop Data Migration Risk Mitigation Plan	Determine Data Security and Privacy Requirements	Determine Data Migration Technology	Determine Data Integration Plan	
Develop Data Migration Communications Plan			Plan Parallel Operation	
Identify Critical Success Factors			Develop Migration Data Quality Plan	
			Develop Data Archival Strategy	
			Develop Data Migration Test Plan	
DATA MIGRATION PLANNING ARTIFACTS				
DATA MIGRATION PLANNING CHECKLIST				

Table 3-2: Data Migration Planning

3.2.3. Planning Data Migration Project

To ensure that both the data migration project and the larger development project are successful, it is good practice to develop a dedicated *Data Migration Project Management Plan* as a subset of the overall *Project Plan*.

Much of the detailed information required by the migration plan is collected during the project, and therefore cannot be included in the initial project plan. This presents two specific issues:

A combined *Project/Data Migration Project Management Plan* would need to include additional contingencies based on forecasts of the results of the initial planning steps.

Alternatively, the combined plan would need to be revised to include strategic and tactical information once migration planning is complete. This would be the same as creating a separate Data Migration Plan at the appropriate time.

For this reason, best practices recommend that a data migration plan that is distinct and separate from the project plan be prepared specifically for the data migration effort.

Once a plan is drafted, the checklist provided in Appendix C can be used to ensure that all relevant points are addressed. When a step is omitted, a simple justification should be given. The sequence of the steps is a suggestion that needs to be reviewed and adapted for each individual migration. In some cases, steps can be performed in parallel; in other cases, steps must be performed in a particular order.

3.2.3.1. Establish Scope

Typically, the scope of a data migration project includes:

- The planning and execution of the transfer of data between a legacy (or source) data storage and a new (or target) storage
- The design of the supporting structures and functions
- The procedures for validating the results along the way

However, a migration project may also include identification of the data source(s) relevant to the migration effort. This provides valuable insight about the level of effort, the timeline, and the resources needed to accomplish the task. Such information also helps identify dependencies and potential risks.

Clear definition of scope at the outset of data migration is important to prevent “mission creep,” which might reduce the project’s chances of success.

3.2.3.2. Identify Risk / Constraints / Dependencies / Assumptions

Each data migration project entails potential obstacles: specifically, risks, constraints, challenges, dependencies, and assumptions. While planning the stages of the data migration effort, identify as many of the obstacles as possible.

Every data migration presents a unique set of issues and risks that must be monitored. Table 2-4 below shows a preliminary list of typical assumptions/dependencies, constraints, and risks to consider:

ASSUMPTIONS/DEPENDENCIES	CONSTRAINTS	RISKS
Sufficient resources are available for all aspects of data migration.	Time/Schedule dictates what must be completed and when.	Unexpected delay and/or downtime might occur.
Sufficient expertise is available for all aspects of data migration.	Funding might limit access to resources that can be devoted to the effort.	The team might encounter complex: <ol style="list-style-type: none"> i. Processes ii. Environments iii. Configuration issues related to data volumes
All environments (legacy, staging, and target) are fully documented, available, and accessible as planned during necessary steps of migration.	Personnel and equipment might be limited or unavailable.	

ASSUMPTIONS/DEPENDENCIES	CONSTRAINTS	RISKS
The team has access to: Subject matter experts for current source system and data	Data requirements and definitions might require clarification by subject matter experts.	Misunderstanding or misinterpretation of requirements might result in a flawed design.
Documentation for data models, physical implementation (database), and business rules, and interfaces Future business requirements	Expertise in legacy-storage environment might be limited due to lack of or outdated documentation.	The team might encounter incompatible software, hardware, processes owing to: Multiple Operating Systems (OS) or vendors Format incompatibilities (Database Management System (DBMS) to DBMS, DBMS to Operating System, etc.)
All selected tools and software packages necessary for data migration will be available and implemented for the necessary migration steps as outlined in the overall project schedule. In addition, the necessary licenses will be available to the project team.	Availability of current and target physical storage (lease conditions, support, physical condition of hardware/software, etc.) might be limited.	Expensive overtime might be required to do certain steps during non-business hours to reduce impact on production.
		Unplanned events or conditions might occur (e.g. configuration of target system is delayed due to illness of support staff.)
		There might be problems with physical relocation of hardware or data.
		Legacy data architecture artifacts might be unavailable or incomplete.
		The new data store might be physically incompatible with the legacy location (such as hard drive connection to server or device drivers).
		The schedule might slip owing to slow or delayed acquisition process of migration software or equipment (e.g., server) and delayed availability.

Table 3-3: Assumptions, Constraints, and Risks

3.2.3.3. Develop Data Migration Risk Mitigation Plan

During this data migration planning phase, the Migration Management Team (MMT) must develop a *contingency strategy* (alternate method of accomplishing the same objective) or *risk mitigation plan* (a means of reducing the impact of undesired results) for each anticipated risk that could jeopardize the successful completion of the migration effort. It is important to continue proactive issue management and proactive risk management throughout the lifecycle of the project. Each data migration project deals with different issues and risks, and therefore each requires a Risk Mitigation Plan specific to the scope of the project.

Note that this Risk Mitigation Plan needs to align with the Risk Mitigation Plan of the overall development project.

A successful data migration effort requires the mitigation of issues and risks to the business/ organization. Identifying these challenges (such as dependencies on other teams within Federal Student Aid, minimal

migration expertise within the organization, or insufficient understanding of data and source systems) and opportunities (such as the identification of the most appropriate data migration method) early in the project allows for proper management and less later disruption. In order to decrease risks, project leads should consider the following actions:

- Migrate only the data required to sustain the future application
- Identify and employ the most appropriate migration method to move the required data into the new solution
- Ensure that all data is migrated accurately and completely
- Ensure that the integrity of the migrated data is maintained
- Minimize disruption to the business during transition
- Prepare a detailed inventory of what data and systems architecture exist, and identify any data issues relevant to the conversion during the early phases of the project
- Identify any resource dependencies, such as access to and availability of environments (source, staging, and target system), tools, software licenses, or personnel. Constraints and restrictions of the target system may require the development of complicated data validation procedures to ensure the integrity and quality of the data loaded.
- Identify staff/resources with knowledge of and experience with the source data. This will reduce the risk of undocumented data issues and will allow identification of potential pitfalls and other issues.
- Identify potential challenges and opportunities early on in the project. This allows for proactive management of these challenges and opportunities.

Risk mitigation is an important task that covers the entire project lifecycle. It is also important to consider that issues and risks detected and addressed in the planning and design phase of a project are less costly than those discovered during the implementation phase.

Best practices indicate that the best way to mitigate risk during a data migration is three-fold:

- Employ commercial data migration software (data profiling, ETL, metadata management, etc.)
- Educate management and technical staff about the features and availability of data migration software
- Reduce costly downtime by selecting an online data migration method⁵.

Even with the most thoroughly tested tools and procedures, whether commercial or custom-built, the conditions necessary for a successful data migration may not occur because they are outside of the control of the data migration project. For instance, if the data migration environment is not configured as per the specifications provided by the Technical Migration Team (TMT), if the team is unable to execute the extract and transformation procedures in this environment, or if there is an insufficient number of software licenses available, the execution of a particular task can only be performed sequentially instead of in parallel. This will prolong the originally-planned duration of this task. These conditions, and the likelihood of their occurrence, should be documented as thoroughly as possible and evaluated through a *Risk Mitigation Matrix*.

This *Risk Mitigation Matrix* should be the standard mechanism for reporting risks and their corresponding contingencies or mitigation solutions. Each identified risk should be given a probability rating and impact level, and a brief statement should be made of a potential mitigation solution. Table 2-5 shows probability ratings.

PROBABILITY RATING	IMPACT LEVEL
High	Likelihood > 70%
Medium	40% < Likelihood < 70%
Low	5 % < Likelihood < 40%

Table 3-4: Data Migration Risk Probability and Impact Levels

⁵ *The Hidden Costs of Data Migration*, page 6

Impact levels are:

- Catastrophic: failure of mission-essential services
- Critical: significantly degraded project performance
- Marginal: significantly degraded support function or secondary mission
- Negligible: inconvenience

Table 2-6 shows selected risks and risk-mitigation strategies, including their probability ratings and impact levels.

RISK DESCRIPTION	PROBABILITY RATING	IMPACT LEVEL	MITIGATION STRATEGY
A Physical Data Model (PDM) does not exist for each legacy system. As a result, the precise structures are unknown to the implementation team, and the schedule is delayed while the missing PDM(s) is/are reverse engineered.	High	Critical	Work closely with Federal Student Aid and legacy contractors to gather PDM documentation early. Upon identifying any missing PDMs, the required database schema shall be reverse-engineered to provide the necessary PDMs.
Data quality issues are not identified until late in the project, thus causing delays and cost overruns.	Medium	Critical	Quality review sessions will be conducted throughout each release so that data quality issues may be identified early and addressed accordingly.
Necessary database personnel are not available during migration.	Low	Critical	In case access to Federal Student Aid resources is very limited, the contractor should consider hiring a short-term consultant to develop the databases to support the target data.

Table 3-5: Data Migration Risk Mitigation Matrix.

Issues arising during the life cycle of the data migration project need to be reported, documented, and resolved as soon as they arise.

The *Data Migration Risk Management Plan* should be reviewed on a regular basis to ensure appropriate monitoring of risks.

3.2.3.4. Develop Data Migration Communications Plan

The *Data Migration Communications Plan* identifies all data management aspects (what, who, when, where, how, about) of the data migration project to stakeholders, Data Migration Team members, and (if needed) external personnel. This plan outlines the recipient, title of communication, content, format, and schedule of each document prepared and shared as a result of the data migration project. Only communications relevant to the migration effort are discussed. The *Data Migration Communications Plan* should cover the following:

- Status reports (weekly, monthly)
- Deliverables and their distribution list including approval authority
- Escalation procedures
- Data profiling findings

- Data cleansing findings
- Information about migration specific metrics, such as:
 - Data volume in source systems and what this means to the project with respect to throughput capability, time to migrate, etc.
 - Performance for data extract/load procedures
 - Sizing of staging area and target system
- Trial migration execution(s) results, including benchmarks
- Important decisions made or reviews passed throughout the lifecycle of the project, including:
 - Stage gate reviews
 - Approval for final data migration (deployment to production)
- Details about the final migration to target system (deployment) including metrics such as:
 - Data volume migrated (total number of records) and statistics on:
 - Accuracy (correct and flawed records in total number and percentage)
 - Completeness (completed and failed records in total number and percentage)
 - Total time to migrated all records
 - Downtime of production system (start to finish)
 - Throughput
- Post-migration activities, including:
 - Validation activities occurred (e.g. test cases performed and results)
 - Steps to resolve remaining migration issues (fix flawed records, research and verify any lost data, etc.)

These communications, and their content, delivery schedule, and format of delivery (electronic or paper), will be outlined in the *Data Migration Communications Plan*.

3.2.3.5. Critical Success Factors for Data Migration Planning

The MMT, in collaboration with the stakeholders, must identify critical success factors for the data migration project. These success factors are the elements considered crucial in ensuring that the data migration effort attains its objective of thoroughly, cleanly, and efficiently transferring the business data from the legacy data environment to the target data environment.

The effort to determine the Critical Success Factors should begin with identifying the information needs of the management level stakeholder community:

“It [Critical Success Factors] focuses on individual managers and their current information needs, whether factual or opinion information. (McNurlin and Sprague, 2006, p. 147).¹⁵

Best Practices in determining critical success factors while planning a data migration effort are, but are not limited to:

- Understand source and target data requirements and structures.
- Define project roles and responsibilities.
- Provide a comprehensive overview and accurate insight into data content, quality, and structure.⁶

⁶ A *Strategic Approach to Data Migration*, page 1

- Document and discuss any anticipated issues or risks with stakeholders and/or business owners.
- Perform migration as an independent project.⁷
- Establish and manage expectations throughout the process.
- Understand current and future data and business requirements.
- Identify individuals with expertise regarding legacy data.⁸
- Collect available documentation regarding legacy system(s).
- Clearly define data migration project roles & responsibilities⁹.
- Prepare a comprehensive overview of data content, data quality, and data structure.¹⁰
- Determine the importance of business data and data quality with business owners and stakeholders.

3.2.4. Determine Data Migration Requirements

Requirements for data migration projects are as varied as they are critical. Outline the business requirements for the data will help determine the data to migrate.¹¹ These requirements may take the form of any necessary agreements, expectations, and/or objectives of the migration.¹² All information is captured in the *Data Migration Requirements* document.

The legacy environment is generally static during a migration, so the Assess Current Environment step (**Section 2.2.5**) should encompass the operational/technical requirements of the current environment. Any requirements for synchronizing changes in content or structure of the legacy environment during the migration must be defined.¹³ The Technical Lead of the MMT must describe in detail any operational/technical requirements for the target and interim (staging area) environments.

3.2.4.1. Determine Business Requirements and Expectations

The MMT consults the business-area stakeholders and subject matter experts (SMEs) regarding any requirements they might impose above and beyond the technical requirements for the data migration.

Together with the business-area stakeholders, the MMT must:

- Establish requirements to be supported by the structural and procedural designs, including:
 - Iterative or phased approach
 - Standard format for all artifacts (exposing metadata for validation (mappings and data element dictionary))
 - Migration/replication design requirements
 - Volume of data

⁷ *Microsoft CRM Data Migration Framework*, page 6

⁸ *Microsoft CRM Data Migration Framework*, page 8

⁹ *Microsoft CRM Data Migration Framework*, page 7

¹⁰ *Strategic Approach to Data Migration*, page 3

¹¹ *Microsoft CRM Data Migration Framework*, page 6

¹² *How to Plan for Data Migration (ComputerWorld, May 21, 2004)*

¹⁵ *Information Systems management in Practice*

¹³ *Taking the Pain Out of Data Migration*, page 1

- Physical relocation of data storage during or after migration
- Required application performance before, during, and after migration
- Establish desired timelines for the individual elements of the data migration, including:
 - Schedule
 - Availability of current and target physical storage (lease-related, support, condition)
 - Hardware availability
 - Allowable downtime during migration
- Identify all business processes that will be used and/or affected by the proposed data migration
- Consider future requirements (build for growth and scalability)
- Determine stakeholder requirements
- Identify relevant business rules and processes
- Identify existing data migration related artifacts for the current environment and approve creation of necessary artifacts where shortcomings are recognized. A technical team will be required if technical artifacts must be created.
- Review Federal Student Aid data security policies and determine applicable security requirements for the data migration

The MMT must also coordinate with any anticipated user community, such as users of the affected software and/or hardware, in addition to other users of the migration technology. The purpose of such coordination is to identify expectations users may have about the effect of the data migration before, during, and after execution.

3.2.4.2. Determine Technology and IT Infrastructure Requirements

Nearly every aspect of a data migration effort can be automated to some degree. It is critical for all of the pieces of a data migration to work together in order successfully to move, cleanse, and/or convert the legacy data to a new environment. Therefore, the technology used at each point in the process should be described in the *Data Migration Plan*.

Together with the business-area stakeholders, the MMT must establish technology requirements. This means defining the conditions and objectives to be satisfied by whatever technology is eventually chosen. A clear understanding of the migration effort will help determine the best technology to use for the migration.¹⁴

For example, ETL software might execute slightly differently against an Oracle database under a UNIX operating system than it does against a Microsoft SQL Server database under a Windows operating system. Such software might not be compatible with the architecture of the staging area. For these and many other reasons, a description of the technology involved in any migration should be prepared.

3.2.4.3. Identify Relevant Best Practices

Best Practices often include an activity that is technically outside the scope of the actual data migration. Considering the requirements of the target data store with regard to longevity and future activity can significantly affect decisions made about and during the migration. It is generally reasonable for the MMT to consider future requirements of the target data store such as durability, migratability, reusability, scalability, future anticipated capacity Determine Data Security and Privacy Requirements.

¹⁴ 2006 *Best Practices for Data Migration*, page 7

The MMT must review and follow the processes and roadmaps outlined in the Handbook for Information Assurance Security Policy Information Assurance Program March 31, 2006 for protection of all data at each source, as well as during the migration of the data between sources.

3.2.5. Assess Current Environment

The assessment of the current environment requires the compilation of all identified artifacts to create a blueprint of the current (legacy) data architecture. In the event that the inventory of data architecture artifacts is incomplete and/or insufficient to address the entire legacy environment, a technical team must complete the architecture.

3.2.5.1. Identify and Collect Existing Data related Architecture

The outcome of this step is the foundation for the data and systems architecture of the data migration project. Documenting and analyzing the current environment from a functional and technical perspective is important for a full understanding of the data and the related business rules and processes. This inventory of facts drives the development of the data migration procedures. Following is a list of sample artifacts that need to be identified and collected:

- Information/data architecture and system architecture documentation of source systems, such as:
 - Logical and physical data models (entity-relationship diagrams and/or repository information of the data structures)
 - Database definition language (DDL) for existing relational databases
 - Data dictionaries documenting each data element (labels and definitions as well as properties)
 - Relevant business rules and processes in the current environment and for the future target system
 - Data mapping from source system to data warehouse
 - Names of systems interfacing with the source systems indicating whether the application sends data to or extracts data from the source system (for context only). This information will help gauge the time constraints for potential downtime of the source system. If there are many interfacing systems, the coordination task will be more complex than with fewer interfaces
 - Data profiling analysis for source system (if available)

Information about known technical constraints of the source system that affected implementation decisions (e.g. limitation on throughput or performance of the server or software version that did) The information collected during the previous task allows for the preparation of a blueprint of the current state of the data architecture for each source system. This blueprint focuses on the logical data model and the data dictionary, if available. If documentation exists only for the physical implementation, this information must be reverse engineered, resulting in a logical data model. The logical data model presents the relationships and the business context of the information used in the application. The physical data model, by contrast, demonstrates the implementation of the data from a technical perspective, meeting performance and data access path requirements. This data provides essential input to the success of the data migration project.

- Not support originally-planned technical solution)
- Information about any known issues or concerns regarding the quality of the available documentation, such as whether the documentation was outdated (e.g. documentation was prepared when the original project started 5 years ago). This information can only be collected through interviews

- Information about any known and identified gaps/missing information that should be resolved after the data migration, which may or may not be documented as a business or technical requirement. This information can only be collected through interviews

3.2.5.2. Blueprint Current State of the Data Architecture

The information collected during the previous task allows for the preparation of a blueprint of the current state of the data architecture for each source system. This blueprint focuses on the logical data model and the data dictionary, if available. If documentation exists only for the physical implementation, this information must be reverse engineered, resulting in a logical data model. The logical data model presents the relationships and the business context of the information used in the application. The physical data model, by contrast, demonstrates the implementation of the data from a technical perspective, meeting performance and data access path requirements. This data provides essential input to the success of the data migration project.

3.2.5.3. Determine Data Migration Technology

Available resources and the migration method(s) selected for the effort will determine most of the technology used for the data migration. As a first step, the technology established by Federal Student Aid, such as the architectural design tools and metadata repositories, should be evaluated to determine whether they meet the requirements. Next, a report stating the shortcomings and the corresponding requirements should be prepared and presented to Federal Student Aid for a decision on how to proceed, if the current technology cannot support all of the requirements.

The MMT will work closely with the DMT to identify the IT infrastructure required to implement the data migration efforts, and any infrastructure affected by the execution of the proposed data migration.

Existing technology, such as the source-data storage devices and software, are already in place, requiring no decision, but rather must be captured as part of the Legacy (or baseline¹⁵) Data Architecture. However, the tools in place to access the legacy data during the migration (such as ETL software or custom programming¹⁶) should be evaluated to ensure migration requirements are met. All technical aspects of the staging area (if used), target environment, actual movement, and validation of the data must be defined and documented by the TMT.

As part of this analysis, the DMT will determine whether the IT infrastructure in place will support the planned data migration effort and, if not, what solution to recommend resolving any shortcomings.

3.2.6. Develop Data Migration Plan

As with any IT project, communication is critical to success¹⁷. The MMT must compile the results of all planning steps and draft the *Data Migration Plan*. The plan shall be submitted to the EDS Team for evaluation. Once the plan has been evaluated for consistency and compliance with enterprise expectations, the plan should be presented to all affected business area owners and stakeholders for feedback and approval.

After all revisions have been finalized and approved, the plan should be provided to the TMT for implementation.

3.2.6.1. Determine Data Migration Method

There are six basic methods of migrating data. These are generally divided into two broad categories based on whether the procedures may be performed while the application remains operational (online) or

¹⁵ *A Practical Guide to Federal Enterprise Architecture, February 2001 (CIO Council), page 5*

¹⁶ *The Complete Data Migration Methodology, page 6*

¹⁷ *How to Plan for Data Migration (ComputerWorld, May 21, 2004)*

the application must be taken out of service (offline) during the actual migration. The basic methods are¹⁸:

- Offline: back up & restore; restore from backup tapes; ftp transfer, and
- Online: array-based replication; volume management or replication; and host-based mirroring

In many cases, a hybrid of these methods is required to satisfy the requirements of a major migration effort. The Data Migration Project Manager, in close collaboration with the Federal Student Aid EDS Team, must determine the method and tools to be used to perform the activities of the data migration. The method can differ based on the legacy systems involved. The method and tools chosen, and the factors contributing to the determination, should be included in the Data Migration Plan. Such factors may include (but are not limited to):

- Distribution (location) of data stores
- Funding constraints
- Available expertise in current and target storage environment (e.g., whether planning is limited to specific options simply because of available expertise)
- Performance (qualitative and quantitative) of procedures/tools
- Source data protection/recovery
- Homogeneous versus heterogeneous storage requirements
- Multi-vendor environment
- Dependencies on external business partners
- Allowable downtime
- Time (schedule) constraints
- Volume of data
- Personnel constraints (availability)
- Complexity of storage and processing environment
- Physical re-location
- Data storage format incompatibilities (DBMS/DBMS, DBMS/OS)
- Configuration issues related to data volume

3.2.6.2. Determine Data Conversion Plan

Migration requirements may require a change to the legacy data during the migration process. There may be changes to form, value, or volume. A strategic approach to data migration that analyzes legacy data at the source will mitigate this risk by allowing analysis both at the source and at each step of the migration process. The MMT, and specifically the data stewards, must define the form and business function of the target data. Transformation of the data values, constraints, and/or format occurs during the migration process, through thoroughly tested rules and procedures.

3.2.6.3. Determine Data Integration Plan

Data migration may require drawing data from more than one legacy data source. The *Integration Plan* describes how conflicts and duplication in source data and data structures will be resolved. The plan also determines how to move the data from the source system(s) to the target system. There are two options:

- Load the data sources sequentially in to the staging area until all source data has been loaded. Then, perform the integration of all source data in the staging area. Finally, move the integrated data to the target data store.
- In some cases, the volume of data or time restrictions may not support the above-described option, and may result in sequential individual data migrations (one for each source system). The staging area could serve as an integration environment to simulate loading the new data set into an environment already populated with operational data.

¹⁸ *Simplifying Technology Refresh with Data Migration Software*, page 6

Staging areas are an optional interim data source, which can serve the purpose of mirroring the ultimate target system. Best practices demonstrate the benefits of establishing a staging area. It allows validation, cleansing, and/or conversion of the integrated data prior to movement into the target location. These trial migrations can be repeated multiple times until the data migration procedures are perfected without affecting the configuration and readiness of the final target system.

The need to integrate multiple legacy-data sources mandates such a staging area.

3.2.6.4. Plan Parallel Operation

Migrating financial systems often require the old and the new system to run in parallel for a pre-defined period of time to ensure the reliability and accuracy of the newly implemented target system. Federal Student Aid follows this principle when planning whether the legacy systems should continue operation for a set time after a successful migration. The legacy system may even serve as a long-term data source for the target system (which is often the case when migrating data from an operational, or transactional, system). However, some legacy systems may be scheduled for complete shutdown upon successful migration of the data to a target system. Others may already be out of operation, which may be the leading factor facilitating the migration.

While different purposes are served by shutting down or continuing operation of legacy systems, the two scenarios have one issue in common: both require that the data contained in the source *and* target systems remain synchronized to some degree as long as both systems are in operation. In the case of a transactional system feeding a data warehouse, the source data is often derived and/or aggregated over a particular time period when being moved into the warehouse. These rules and algorithms shall be developed as part of the procedures for populating the target data store.

A third scenario involves maintaining the legacy data store for a period of time while the operation of the new system is validated. While this is generally considered a post-migration task, the full operation of the new system may reveal errors in the migrated data, requiring revisions to some part of the data migration (data quality remediation, data migration procedures, etc.). Roadmaps for operating the two systems in parallel, monitoring and comparing the performance of each system, and resolving issues as they arise should be established and included in the *Data Migration Plan*.

3.2.6.5. Develop Migration Data Quality Plan

The *Data Quality Plan* concentrates on the quality of the legacy data. It requires multiple efforts that can be performed in parallel. All outcomes will determine the overall data quality of the source system(s) to be migrated. In addition, the *Data Quality Metrics* and *Data Loss Tolerance* information will be used as benchmarks to determine the fitness of the data for deployment.

Define Data Quality Metrics: A *Proof of Concept*, which simulates a full data migration by operating on a sampling of data supporting a single event, such as a single transaction or single concept¹⁹ may be performed if a commercial data migration (such as ETL software) or data profiling tool is used. The *Proof of Concept* provides a field-level and/or record-level view of the legacy data and helps identify anomalies. The document will validate the compatibility of the technology selected to perform the data migration, and will provide data quality metrics based on the sample data that may be used to project the level of effort required to perform full data remediation. If custom software or procedures are planned for the data migration, then data metrics must still be established for measuring the quality and integrity of the data before, during, and after each migration stage.

Define Data Loss Tolerance: If all data stores that participate in a data migration effort (legacy, staging, and target) have the same basic specifications (e.g., a relational database using version X of RDBMS Y on operating system Z, etc.), it is reasonable to expect that all data will transfer without loss. However, on occasion obsolete data structure or formats may not translate 100% into a modern environment. In this scenario, the MMT must consult the business stakeholders to determine the tolerance level for data loss.

Establish Data Quality Remediation Plan: The creation of a “zero-defect” data quality policy is optimal prior to data migration. Such a policy can be put into place by performing error correction, including

¹⁹ *A Strategic Approach to Data Migration*, page 3

passive remediation (at the source) or, through active remediation, which corrects data errors *during* the migration process. If this cannot be accomplished, then fixing known or discovered errors in the legacy data should be the first post-migration step²⁰. Once metrics are established, the MMT must determine at what stage of the migration, and by what means, the identified data quality issues shall be remedied, and lay these decisions out in the *Data Quality Remediation Plan*.

If a passive remediation plan is chosen that affects the content of the source (legacy) data, notification to dependent systems of changes must be included in the *Communications Plan*.

3.2.6.6. Develop Data Archival Strategy

The plan for managing data once it is no longer necessary for immediate access and use is called the *Data Archival Strategy*. The MMT must interview stakeholders and formulate strategies regarding what data to retain, how long, where and how. A thorough architecture of the legacy system may already include a *Data Archival Strategy*, but it likely only covers the retention and management of data during the operational life of the legacy system.

If the strategy in place for the legacy system is sufficient to address the retention of data once the system is removed from operation, then the full strategy may be adopted as part of the *Data Archival Strategy* within the *Data Migration Plan*. If, however, the strategy does not address system shutdown or does not exist at all, then the MMT must establish a strategy for retaining and retrieving the legacy data after the legacy system is taken out of operation.

3.2.6.7. Develop Data Migration Test Plan

The MMT must establish a plan for testing the migration procedures at each step. All data movement procedures, transformation/conversion procedures, data cleansing procedures, and data validation procedures must be accounted for in the context of the *Migration Data Architecture*. The data migration procedures must be able to successfully satisfy the requirements set forth in the data requirements (*as demonstrated in the test plan*) before proceeding to the full migration.

3.2.7. Define and Assign Roles & Responsibilities

3.2.7.1. Define Migration Roles and Responsibilities

It is recommended that specific roles and responsibilities be established early on during the planning stages to help guide the DMT throughout the project²¹.

Table 2-7 shows the data migration project's specific roles and responsibilities for stakeholders and team members.

DATA MIGRATION ROLES	RESPONSIBILITIES
Stakeholders (Federal Student Aid Business Owners, Enterprise Architects, etc.)	Provide guidance and input to the overall project Determine acceptance and success criteria for data migration Approve the Data Migration Plan and other artifacts prepared by the data migration project team Approve production-readiness of migration procedures and the timeline for deployment of data migration.
Enterprise Data Services (EDS) Migration Team (May include Program Manager, Data Architects, Governance Team members, etc.)	Provide guidance to Business Area data migration teams in establishing Data Migration Plans Evaluate the Data Migration Plan prior to approval by the Stakeholders and implementation by the Technical Migration Team Provide the Technical Migration Team with any standard enterprise metadata (standards, policies & procedures, designs, etc.) necessary to develop the Data Migration Architecture

²⁰ *A Strategic Approach to Data Migration*, page 1

²¹ *Microsoft CRM Data Migration Framework*, page 7

DATA MIGRATION ROLES	RESPONSIBILITIES
Migration Management Team (MMT)	Works under the auspices of the Migration Manager to plan and manage the Data Migration project
Project Manager ²²	Plan the implementation effort (of which the migration project is a subset or stream) Execute the implementation project plan Monitor scheduling, progress, performance, and issues and risks Close the project
Migration Manager (Could be the same as Project Manager based on scope/budget of the project)	Prepare the Data Migration Plan (a subset or aspect of the overall project implementation plan) in collaboration with Project Manager and data migration project team Manage and monitor the Implementation & Validation of the data migration Collaborate with the Project Manager on risk mitigation Document data migration results Report to the Project Manager
Data Steward(s)	Planning: Compile legacy data architecture, including models, data dictionaries, volume metrics, and other artifacts Analysis & Design: Coordinate / validate data designs and inter-environment (source / staging / target) correlations; facilitate development and validation of data profile metrics Implementation: Coordinate data cleansing and data quality remediation Closeout: Coordinate validation of reports on the results of the data migration
Migration Technical Lead (MTL)	Provide technical expertise to the Migration Management Team during Planning Compile and document technical requirements during Planning and Analysis & Design Lead the Technical Migration Team during Analysis & Design, Implementation, and Closeout Provide technical results and statistics of the migration to the Migration Manager during Closeout
Technical Migration Team (TMT) ²³ (May include Data Architects, Business Analysts, Database Administrators, Programmers, Technical Writers, & and experts in various technology technologies employed during migration)	Contribute technical expertise during Planning Document legacy data architecture during Planning, if necessary Analyze & Design Migration Data Architecture (staging & target) Execute the Data Migration Plan Develop migration procedures and perform trial migrations; refine procedures as needed Validate data migration results Compile technical results of migration on behalf of the Migration Technical Lead

Table 3-6: Data Migration Roles and Responsibilities

3.2.8. Data Migration Planning Deliverables

Artifacts of the planning phase include:

- A dedicated *Project Management Plan*
- A *Data Migration Plan (as a component of the overall Project Management Plan)*
- A *Data Migration Requirements* document

²² PMP In Depth, page 8

²³ 2006 *Best Practices for Data Migration*, page 6

- A *Risk Mitigation Plan*
- A *Risk Mitigation Matrix*
- A *Consolidated Legacy Entity Relationship Diagram (ERD)*
- A *Communications Plan*
- A *Data Conversion Plan*
- A *Data Integration Plan*
- A *Parallel Operation Plan*
- A *Data Migration Quality Plan*
- A *Data Migration Data Quality Remediation Plan*
- A *Data Migration Archival Strategy*
- A *Data Migration Test Plan*
- A *Data Migration Roles & Responsibilities* document

3.2.9. Data Migration Planning Checklist

The MMT may use the checklist provided in Appendix C to ensure that all aspects of planning set forth in this methodology are accounted for in the *Data Migration Plan*.

3.3. Data Migration Analysis and Design

3.3.1. Analysis and Design Overview

A thorough *Data Migration Plan* shall include either requirements for the Analysis and Design activities or justifications for not including a particular activity.

From the Legacy (or source) Data Architecture, the completed analysis, and the *Data Migration Plan*, the TMT must design the Migration Data Architecture. This Architecture is comprised of:

- The ‘as is’ legacy architecture (blueprint)
- The ‘to be’ Staging Area Data Architecture, if necessary
- The ‘to be’ Target Data Architecture
- The correlations (‘mappings’) showing inter-architecture relationships and instructions to turn data from one data architecture into valid data in the next (source into staging or target, staging into target)

The data structures and data migration procedures must satisfy the data migration requirements. The resulting artifacts consist of fully-attributed logical data models, data dictionaries, function (or process) models, mappings between the “as is” and “to be” data structures with corresponding business rules and transformation logic, and any other applicable artifact defined in the *Federal Student Aid Integrated Architecture Framework*. These artifacts should then be vetted through the EDS Team and business area stakeholders.

3.3.2. Data Migration Analysis and Design Tasks and Subtasks

The following table presents the four major tasks of the Data Migration Analysis & Design Phase. Subtasks may occur in parallel or in the sequence shown in this chart. The Data Migration Project Manager, in collaboration with the DMT, will determine the order in which these tasks will be performed.

PERFORM DATA MIGRATION ANALYSIS	DETERMINE DATA SECURITY CONTROLS	DESIGN DATA MIGRATION ENVIRONMENTS	DESIGN DATA MIGRATION PROCEDURES
Analyze Current Environment	Determine Enterprise Management and Operational Security Controls	Design Staging Area	Design Data Staging Procedures
Evaluate Data Migration Technology		Design Target Data Architecture	Design Data Cleansing Procedures
Evaluate Data Quality		Correlate Migration Data (Source/Staging/Target)	Design Data Conversion Procedures
Perform Data Profiling		Determine Data Migration Technology Configuration	Design Target Data Migration Procedures
Critical Success Factors for Data Migration Analysis and Design			Design Data Validation Procedures
			Design Data Quality Remediation Procedures.
			Refine Data Migration Test Plan

Table 3-7: Data Migration Analysis and Design

3.3.3. Perform Data Migration Analysis

3.3.3.1. Analyze Current Environment

The MMT, and specifically the Technical Lead, must fully understand the IT environment in place and any constraints or technical limitations. A thorough review of review all compiled artifacts related to the legacy environment is suggested. It is important to determine whether current (or new) limitations and/or constraints will exist in the new target environment. This information will affect the design of the data migration activities and procedures.

3.3.3.2. Evaluate Data Migration Technology

The MMT, and specifically the Technical Lead, must review all available technology options. Should the technology in place not meet the requirements, the team should select and recommend tools best suited to satisfying the requirements of the data migration.

Data Migration infrastructure should be in place during the design phase of the project.

3.3.3.3. Evaluate Data Quality

The MMT, and specifically the Data Steward, must evaluate the current state of the legacy data in accordance with the *Data Migration Plan*, and specifically the *Data Quality Metrics*.

The MMT, and the Data Steward, determine how to assess the quality of the legacy (source) data (*Data Quality Metrics*) and how to remedy identified quality issues in collaboration with the business owners and stakeholders. The actual remediation steps will be defined after the data analysis is performed, and after and the Data Quality Metrics are applied. These actions result in the refined *Data Quality Remediation Plan*.

There are two types of potential data quality issues, each of which requires a different set of remediation activities:

- Non-compliance with the metadata structure (social security number data field containing letters instead of numbers).
- Bad or corrupted data in the database (incomplete social security number; multiple records for the “same” person).

The procedures for correcting any data quality issues shall be developed and executed during the Implementation stage of the migration. Data Quality measures the compliance with business and data validation rules implemented by the analyzed application. Additional analysis can be done by spot-checking with custom-build SQL-queries for bad data such as corrupted records. The results of the data analysis determine the need for data cleansing, and the data quality metrics will indicate what specific action needs to be taken.

3.3.3.4. Perform Data Profiling

Data Profiling²⁴ is the initial assessment of the legacy data (structure and content) to understand and determine any quality challenges. There are two types of Data Profiling:

- Metadata profiling: assessment and examination of the data structures in place. The results from this activity can be used to evaluate compliance with enterprise-wide standards.
- Content profiling: assessment and examination of the data content. The results of this assessment reflect the quality of the content of the data captured, and identify issues that will be resolved through data cleansing.

The assessment is a process whereby the team examines the data available in an existing database and collects statistics and information about that data. The purpose of these statistics is to:

- Give metrics on data quality, including whether the data conforms to company standards
- Assess the risk involved in integrating data for new applications,
- Monitor and track data quality
- Assess whether metadata accurately describes the actual values in the source database

Profiling activities should follow the following three steps in order presented:

- **Step 1 — Column Profiling:** Provides critical metadata.
- **Step 2 — Dependency Profiling:** Identifies intra-table dependencies. Dependency profiling relates to the normalization of a data source, and addresses whether or not there are non-key attributes that determine or are dependent on other non-key attributes. The existence of transitive dependencies here might be evidence of second-normal form
- **Step 3 — Redundancy Profiling:** Identifies overlapping values between tables. This is typically used to identify candidate foreign keys within tables, to validate attributes that should

²⁴ Wikipedia: [Wikipedia- Data Profiling](#)

be foreign keys (but that might not have constraints to enforce integrity), and to identify other areas of data redundancy. Example: redundancy analysis could provide the analyst with the fact that 80% of the time, the ZIP field in table A contained the same values as the ZIP_CODE field in table B.

Column profiling provides critical metadata, which is required in order to perform dependency profiling, and as such must be executed before dependency profiling. Similarly, dependency profiling must be performed before redundancy profiling.

The use of automated data profiling tools is a Best Practice in the data profiling step in data migration²⁵.

This step partially overlaps with blueprinting the state of the legacy architecture. Automated profiling tools might be used to facilitate the procedures, but the procedures might also be done manually in the absence of automated software. Once completed, a Data Profile contributes to Data Conversion and Data Quality Remediation. The documented results of the analysis become a resource for the design of the staging and target architectures. The *Data Profile Assessment* needs to be distributed and presented for approval as outlined in the Communications Plan to discuss and determine the criticality of the findings.

In addition, Federal Student Aid must analyze which of the identified Data Quality issues can be resolved through Data Cleansing. The questions below help in determine the most appropriate Data Cleansing approach and responsibilities:

- Where do the identified data issues originate?

There are two possibilities: They were introduced by Federal Student Aid applications or through data received from their business partners as part of the data exchange.

- How can the identified data issues be repaired?

It might be possible to implement strong validation rules at the front end (GUI) and prevent the entry of invalid data at the point of data entry; or validation rules can be implemented at the database level.

- Do the data issues refer to historical data only (e.g. data older than 5-10 years)?

It is possible that improved business and validation rules have been implemented after detection of these data issues and that newer data is in better condition. A decision needs to be made as to whether it is worthwhile and/or necessary to repair the historical data records.

- Are the identified data issues caused by missing validation rules?

Design and implementation of proper business rules could repair these issues.

- Who owns the data and is responsible for the data quality?

Usually, the business owners are responsible to ensure high data quality from a business perspective. The business owners, in collaboration with the Data Steward(s), should define the necessary steps for data cleansing and prepare a plan and timeline for implementation.

3.3.3.5. Critical Success Factors for Data Migration Analysis & Design

Best practices recommend establishing the following goals while designing target data structures and data migration procedures:

- Understand data requirements (architecture and business rules),
- Design comprehensive data migration procedures upon understanding of data content, quality, and structure²⁶, and
- Leverage standardized data structures (Enterprise Data Architecture).

²⁵ *A Strategic Approach to Data Migration*, page 2

²⁶ *Strategic Approach to Data Migration*, page 3

3.3.4. Determine Data Security Controls

3.3.4.1. Determine Enterprise Management and Operational Security Controls

Each data migration project needs to operate within the security controls in place at Federal Student Aid. Therefore, the MMT must work with Risk Management to ensure the appropriate processes outlined in the “*Handbook for Information Assurance Security Policy Information Assurance Program March 31, 2006*” are followed to protect all data at each source and during the migration of the data between sources. This guide also addresses the following operational security controls of:

- Personnel security
- Physical and environmental protection
- Production controls
- Contingency planning
- System hardware controls
- Software maintenance controls
- Data integrity/validation controls
- Documentation standards
- Security awareness training
- Incident reporting
- Data encryption

3.3.5. Design Data Migration Environments

3.3.5.1. Design Staging Area

The data migration transfers data from one or more legacy (source) data stores to a target data store. The migration is seldom a *direct* transfer between the source and target. Staging the data in an interim location allows for additional processing. Most of the activities created during planning are best performed in a staging area, specifically Data Conversion / Transformation and Data Quality Remediation.

It is possible (and at times preferable) to perform quality remediation (fixing errors) within the actual legacy (source) data location, but it is *not* recommended to perform quality remediation within the target data location, especially if the target is already an operational data location.

If either of these activities requires a custom data area, then the TMT must design the Staging Area Data Architecture. The data structures must house the interim data during the (trial) migration, and the procedures must populate the structures throughout the migration(s). To ensure the success and validity of the data migration procedures and data architecture, it is essential that the staging environment mirror the production environment. Otherwise, the results of the trial migrations might not be representative.

Another strong reason for processing the migrating data within a staging area is the need to reconcile or integrate data coming from separate data sources. If data representing the same information is structured differently in different sources, a staging area provides the means to consolidate all of the source data and validate it before transferring the data to the target data location. Also, if data is being migrated in phases — different sources at different times or only a part of the data at a time — the staging area provides an area to process and transform the migrating data before loading it to the target data location.

The proposed Staging Data Architecture needs to be distributed and presented for approval as outlined in the *Communications Plan*.

3.3.5.2. Design Target Data Architecture

The TMT must establish the Target Data Architecture. The target data structures must house the data to support the target application, and the procedures must populate the target data store from the staging area (or the source data store, if a staging area is not used during the migration). The data structures to house the target data should satisfy the data requirements of the target application or business area.

In addition to the data structures required for the target data store, the TMT must design procedures to:

- Move either the staged data or the source data into the target data store
- Integrate migrated data into the target data store, if the target data store is already populated
- Validate that previously existing data structures and functionality remain intact after integrating migrated data into a target data store with already operational data. This scenario also occurs when performing sequential data migrations due to large volumes into one target data store
- Validate the migrated data for completeness and accuracy

The proposed Target Data Architecture needs to be distributed and presented for approval as outlined in the *Communications Plan*.

3.3.5.3. Correlate Migration Data (Source/Staging/Target)

The TMT must now correlate, or map, the individual architectures to each other, providing a roadmap for the data — and the migration team — to follow from the original data source to the final (target) data source. If an interim staging area is used, then the path should, obviously, pass through the staging area.

In addition, the correlations must also include any mapping of the data to the rules, translations, and/or transformation procedures to which the data must adhere when moving from one data location to another. The *Data Correlation Report* needs to be distributed and presented for approval as outlined in the *Communications Plan*.

3.3.6. Design Data Migration Procedures

3.3.6.1. Design Data Staging Procedures

The TMT designs the necessary procedures to extract the data from the legacy (source) data store, transport the data to the staging area, and perform any necessary translations, or transformations, prior to populating the staging area. Sources of input for this task include:

- Blueprint of the existing data architecture
- Design of the staging area data store
- Data correlation report
- Any identified rules, translations, or transformations that the data must undergo to meet the staging/target data structure requirements.

In cases of data migration projects where a staging area is not used, the data extract and transport procedures are still required. This activity may overlap with the activity of designing target migration procedures. In addition to the design of this task, and of all following design tasks, the team must work with representatives from the business owners and stakeholders to develop a set of test cases to validate the procedures.

3.3.6.2. Design Data Cleansing Procedures

The TMT designs the necessary procedures to fix errors and refine the source data, either within the legacy (source) data store or within the staging area, based on input from the business owners and other stakeholders.

As part of this activity, the ownership and responsibility for the execution of the data cleansing task must be determined. In many cases, when the TMT consists mainly of contracting staff, Federal Student Aid

employees will perform the data cleansing. This situation should be treated as a potential risk with respect to the timely completion of the task.

3.3.6.3. Design Data Conversion Procedures

The TMT designs the necessary procedures to convert the source data to the proper values and formats required in the staging and target data store. Sources of input for this task include:

- Blueprint of the existing data architecture
- Design of the staging area and the target data store
- Data correlation report
- Any identified rules, translations, or transformations that the data must undergo to meet the staging/target data structure requirements

3.3.6.4. Design Target Data Migration Procedures

The TMT designs the necessary procedures to extract the data from the staging or source data stores as appropriate, transport the data to the target data store, and in case no staging area is used, perform any necessary translations or transformations prior to populating the target data store.

3.3.6.5. Design Data Validation Procedures

The TMT designs the necessary procedures to validate the integrity of the data content at each stage of the migration. Validation procedures must also support validation of the completeness and accuracy of the migrated data.

3.3.6.6. Design Data Quality Remediation Procedures

The TMT designs the procedures to be used to remediate identified data quality issues through the data profiling. These procedures will be developed in close collaboration with Federal Student Aid business owners to meet current (and future) business requirements, and in collaboration with the EDS Team with regard to compliance with Federal Student Aid data standards.

3.3.6.7. Refine Data Migration Test Plan

All individual test cases must be consolidated into the overall *Data Migration Test Plan*. The TMT uses this detailed information to update and refine the original plan through sequencing the execution of these test cases and identification of any dependencies.

3.3.7. Data Migration Analysis & Design Deliverables

Artifacts of the planning phase include:

- *A Data Profiling Assessment & Report (Legacy System)*
- *A Data Migration Data Quality Remediation Procedures Design*
- *A Migration Data Architecture* covering:
 - Staging Area and Target Data Architecture Design including
 - *Logical Data Model (Staging and Target Area)*
 - *Physical Data Model (Staging and Target Area)*
 - *Data Dictionaries (Staging and Target Area)*
 - Migration Data Correlation Report
 - Data Migration Technology Configuration
- *A Data Migration Procedures Design*
 - A Data Staging Procedures Design
 - A Data Cleansing Procedures Design

- A Data Conversion Procedures Design
- A Target Data Migration Procedures Design
- A Target Data Validation Procedures Design
- A Data Migration Test Plan (refined)
- A Data Load Plan
- A Data Back-up & Recovery Plan (including back-out)

3.3.8. Design Checklist

The MMT may use the checklist provided in Appendix C to ensure that all activities of Analysis and Design set forth in this methodology are accounted for in the *Data Migration Plan* and the *Migration Data Architecture*. As the *Data Migration Plan* and/or *Migration Data Architecture* are submitted to the stakeholders and EDS Team for approval, the same checklist may be used, with supporting commentary for feedback, to evaluate the proposed architecture and thoroughness of the documentation.

3.4. Implementation

3.4.1. Implementation Overview

The steps of Implementation and Validation are logically interdependent, much like the steps of Planning and Analysis and Design. Although not every activity may occur for every migration, the steps that *do* occur should generally follow the sequence shown below. There is one exception: As discussed in several Best Practice articles, a decision should be made during Planning as to the most efficient and effective time to do Data Cleansing and Data Conversion.

If the data migration covers multiple source systems, repeat the steps of extracting the data and loading it in to the staging area until all source data has been loaded. The staging area supports the integration of all data. Figure 2-5 shows the high-level process flow of the Implementation Phase.

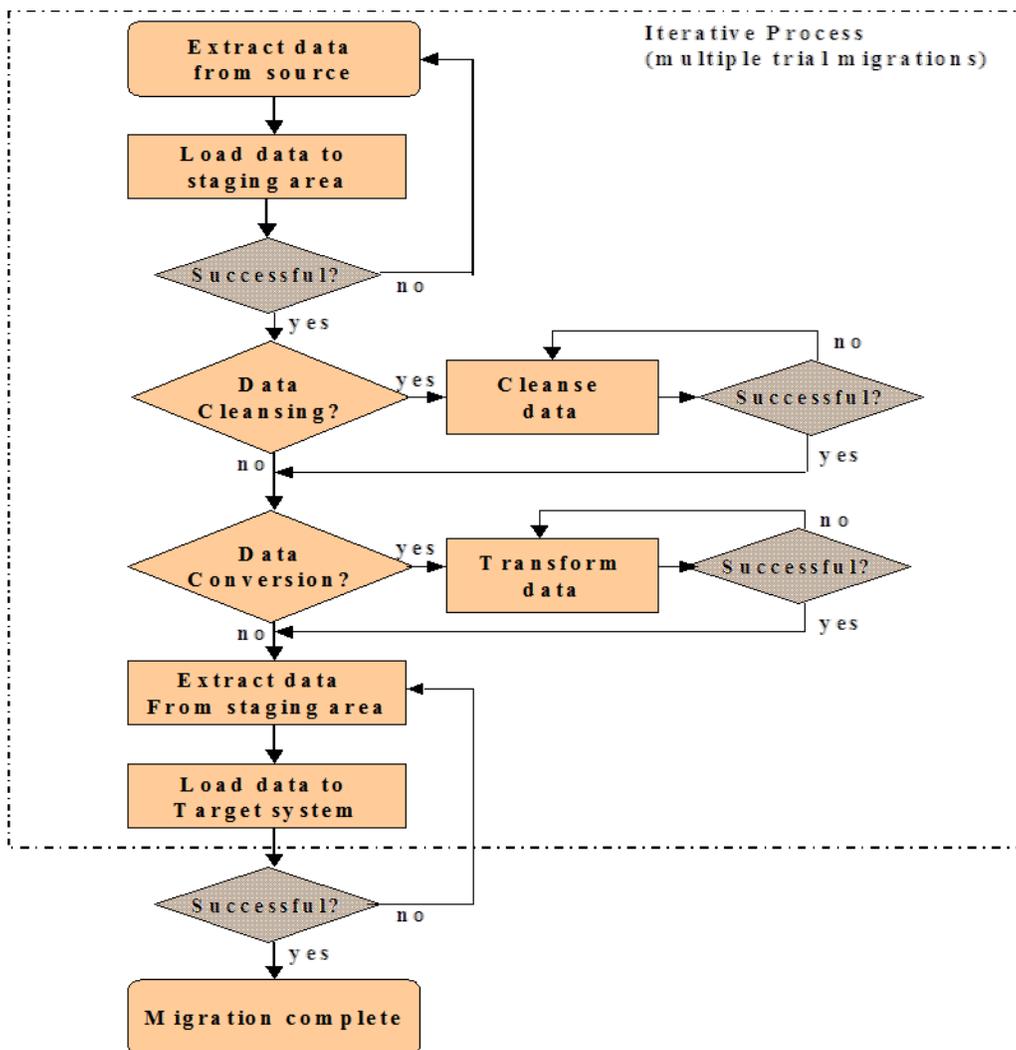


Figure 3-5: Data Migration Implementation Plan Flow

3.4.2. Data Migration Implementation Tasks & Subtasks

The following table presents the six major tasks of the Data Migration Implementation Phase. Subtasks may occur in parallel or in the sequence shown in this chart. The Data Migration Project Manager, in collaboration with the DMT, will determine the order in which these tasks will be performed.

DEVELOP DATA MIGRATION PROCEDURES	STAGE DATA	CLEANSE DATA	CONVERT/ TRANSFORM DATA	MIGRATE DATA	POST-MIGRATION
Configure Resources	Create Staging Area	Cleanse Data according to Data Remediation Plan	Convert/ Transform Data	Perform Trial Migration	Operate Legacy and Target Environment in Parallel
Develop and Test Data Migration Procedures	Populate Staging Area	Validate Cleansed Data	Validate Converted/ Transformed Data	Validate Results of Trial Migration	Validate Parallel Operation
Develop and Test Validation Procedures	Integrate Staged Data			Obtain Approval for Full Migration into Production Environment	Release Data Environments
Develop and Test Data Cleansing Procedures	Validate Staged Data			Perform Full Migration (Deployment)	
Develop and Test Data Conversion/ Transformation Procedures				Validate Results of Full Migration	
Establish Access to Staging and Target Area					
Critical Success Factors for Implementation					

Table 3-8: Data Migration Implementation

Using the *Data Migration Data Quality Plan* (here, using the Data Loss Tolerance and Data Quality Metrics as a benchmark), the TMT must execute the validation procedures to measure the success of each stage of migration, and must also validate any stage of the implementation that changes the data (such as format, content, or location). Any time the resulting data fails the validation; the procedures of that migration stage should reverse, or “roll back,” the data and make any needed corrections. The trial migrations should be repeated until acceptance and/or success criteria are met and the migration procedures and environments are ready for deployment.

3.4.3. Develop Data Migration Procedures

3.4.3.1. Configure Resources

In some cases, data migration planning and implementation are separately-funded tasks. This funding arrangement can put constraints on resource availability. If not done during project start-up, all resources to be used during the migration must be acquired, assigned, installed, and/or configured prior to implementation.

Do the following:

- Assign personnel of the TMT to the various tasks required to migrate the data
- Define Create, Read, Update and Delete (CRUD) Matrix outline access privileges to Staging Area and Target Data Store for team members

- Request or acquire the data access required by the technical team to implement the various procedures (if not already done)

The following technical steps must now occur:

- Establish test environments
- Make test environments available to the technical team
- Install and/or configure all software involved in the migration²⁷

These steps carry a potential risk and dependency because both the creation of the test environment and the configuration and implementation of the software will most likely be performed by staff outside the immediate project.

3.4.3.2. Develop and Test Data Migration Procedures

The TMT develops and tests the data migration procedures in accordance with the *Data Migration Plan*, the *Data Migration Test Plan*, and the documented *Migration Data Architecture*.

3.4.3.3. Develop and Test Data Validation Procedures

The TMT (with input from stakeholders and business users) develops and tests the data validation procedures in accordance with the *Data Migration Plan*, the *Data Migration Test Plan*, and the *Migration Data Architecture*; as well as in compliance with all applicable business rules and processes. These procedures must be capable of confirming that data loss and accuracy is within allowable parameters during each stage of the migration. If 100% migration of data between different environments is not feasible, the MMT must consult with business stakeholders and revise tolerances for data loss and update the data loss tolerance in the *Data Migration Plan* with the revised tolerances. The validation procedures at each stage must take these tolerances into account.

3.4.3.4. Develop and Test Data Cleansing Procedures

The TMT develops and tests the data cleansing procedures in accordance with the *Data Migration Plan*, and the *Data Migration Test Plan*, specifically the *Data Quality Metrics* and *Data Loss Tolerance*. If the procedures designed for data cleansing cannot be scripted or saved (possibly because they are steps to follow in a commercial software package or require significant manual operation), then they should be thoroughly documented and tested at this stage.

3.4.3.5. Develop and Test Data Conversion/Transformation Procedures

The TMT develops the data conversion and/or transformation procedures in accordance with the *Data Migration Plan*, and the *Data Migration Test Plan*, specifically the *Data Conversion Plan*. As mentioned before, if the procedures designed for data conversion cannot be scripted or saved (possibly because they are steps to follow in a commercial software package or require significant manual operation), then they should be thoroughly documented and tested at this stage.

3.4.3.6. Establish Access to Staging and Target Areas

The TMT must coordinate with the proper points of contact at Federal Student Aid to gain all necessary access to the software and establish the user accounts that are required to perform all operations of the data migration. Omitting or delaying this task can result in a delay of the data migration project.

There is a potential risk and dependency because Federal Student Aid staff outside of the immediate project will perform this task.

3.4.3.7. Critical Success Factors for Implementation

Best Practices recommend the following activities based on critical success factors during migration of data from one or multiple source data stores to a target data store:

- Execution of a thorough and detailed Migration Test Plan

²⁷ *Microsoft CRM Data Migration Framework*, page 8

- Expertise in all data migration software²⁸
- Expertise in all COTS and custom software used during the migration
- Comprehensive understanding of Data Migration Plan and Migration Data Architecture²⁹
- Following and compliance with Risk Mitigation Plan
- Following and compliance with change control and quality assurance procedures

3.4.4. Stage Data

3.4.4.1. Create Staging Area

The data staging area can be created manually and/or automated. Either the TMT or Federal Student Aid staff can perform this critical task. Completion of this task is a precondition for using the staging area. Missing or delaying this task can cause a delay of the migration project.

This task also carries a potential priority based risk and dependency in that outside staff will most likely perform it and not managed by the Project Manager.

3.4.4.2. Populate Staging Area

The TMT executes the data migration procedures that transport (and transform, if applicable) data from the source data store to the staging area in accordance with the *Data Migration Plan* and documented *Migration Data Architecture*. If the staging of the data is to be done in phases, then the phases must be performed according to the established schedule.

After completion of the task, the team collects statistics for reporting and validation purposes, as laid out in the *Data Migration Plan*.

3.4.4.3. Integrate Staged Data

This step may not be independent of the steps required to populate the data in the staging area. The importance of *integrating* data from multiple sources must be clearly addressed in any migration plan. Depending on the age and condition of the source data, it might even be necessary to integrate multiple renditions of the same data from a single source, an activity that is commonly referred to as *reconciling* the source data.

3.4.4.4. Validate Staged Data

The TMT executes the validation procedures to measure the success of the staging of the source data. Upon completion, the team should then compare the results to the *Data Quality Metrics* and acceptance criteria set by the stakeholders and disseminate the outcome according to the *Communications Plan*. Also, if the threshold cannot be met after several trial migrations, stakeholders need to approve or deny requests to move v forward (a decision point that is a dependency and risk to the project).

When multiple source systems are being merged into one target system, it is recommended to perform data validation after each load from each source system.

3.4.5. Cleanse Data

3.4.5.1. Cleanse Data According to Data Remediation Plan

The TMT must execute any procedures defined for measuring data quality. The data quality may be evaluated in the source data store, in the staging area, or in both. The TMT should then evaluate the results in the context of the *Data Quality Metrics* from the *Data Migration Plan*. If data errors need to be corrected, the *Data Quality Remediation* should be followed.

²⁸ *Microsoft CRM Data Migration Framework*, page 8

²⁹ *Strategic Approach to Data Migration*, page 3

This step carries both potential risk and dependency, as Federal Student Aid staff most likely performs the data cleansing.

3.4.5.2. Validate Cleansed Data

The TMT must execute the validation procedures to measure the success of the cleansing stage of the data migration

3.4.6. Convert / Transform Data

3.4.6.1. Convert / Transform Data

The TMT must execute any procedures defined for converting the source data format to the format required by the staging and/or target data store. If the data content must undergo transformation, the TMT should execute those procedures during this stage of the migration.

3.4.6.2. Validate Converted / Transformed Data

The TMT must execute the validation procedures to measure the success of the Data Conversion and/or Transformation. With commercial ETL software, these procedures may be built into the actual execution of the transformation. The resulting statistics should then be collected and presented for review.

3.4.7. Migrate Data

3.4.7.1. Perform Trial Data Migration

The TMT must execute the data migration procedures that transport (and transform, if applicable) data from the staging area, or the source data store, to the target data store in accordance with the *Data Migration Plan* and the documented *Migration Data Architecture*. These procedures may be performed on the full set of data or, if planned, on a representative subset. If the target environment is already operational, the target of the trial migration may be a test environment that mirrors the production environment. The results of the trial data migration must be validated immediately upon completion.

If the validation deems the migration 100% successful, and the trial migration was done on the full set of source data, the results may be accepted as the full migration as soon as the business area stakeholders approve them. If errors occur during the trial migration, or if the migration was done on an incomplete set of data or in a test environment, the necessary corrections must be made and the trial migration must be repeated.

If the trial migration was successful on a subset of data or on data directed to a test environment, the team may proceed to do the full data migration.

3.4.7.2. Validate Results of Trial Migration

The TMT must execute the validation procedures to measure the success of the trial migration and collect the results for review.

3.4.7.3. Obtain Approval for Full Migration into Production Environment

Once a successful trial migration has been executed and the results have been reviewed and approved by the business area stakeholders and the EDS Team, the TMT must proceed with the execution of the full data migration. If the final, successful trial migration satisfied all requirements and success criteria, the migration may be released for production.

3.4.7.4. Perform Full Data Migration--Deployment

Upon approval, the TMT will execute the data migration procedures in the production environment in accordance with the *Data Migration Plan*, and documented *Migration Data Architecture*.

If resource limitations or the migration of large volumes of data make it necessary, the full data migration may be done in stages.

3.4.7.5. Validate Results of Full Migration

The TMT must execute the validation procedures to measure the success of the full data migration and collect the results for review.

3.4.8. Post-Migration

3.4.8.1. Operate Legacy & Target Environments in Parallel

Once the full migration has been executed, validated, and approved, the business area stakeholders may desire that the legacy system and the target system operate in parallel for a period of time. This decision, the duration of parallel operation, and the criteria for concluding the parallel operation should be captured in the *Data Migration Plan*, specifically in the *Parallel Operation Plan*.

3.4.8.2. Validate Parallel Operation

During the “trial” period of parallel operation if errors are detected in the operation of the target system then (1) the legacy system may continue to operate according to original requirements, and (2) the data migration may be revisited, repaired, and repeated.

3.4.8.3. Release Data Environments

Upon confirmation that the target environment is operating in a satisfactory manner, the source data environment may be released. If the source environment is to be retired from operations, retirement must follow a system retirement plan and, may occur at this time.

Any staging area, if used, would generally be taken out of operation at this point in accordance with the archival strategy developed as part of the overall *Data Migration Plan*. However, it is important that the environment not be purged entirely, in case errors in the data migration are discovered. It may be possible to restore the staging environment and re-populate the target environment without having to repeat the data migration procedures.

If the staging area is part of a long-term solution – such as an interim data store between an operational data store and a data warehouse – it should be approved and placed into full-time operation. The target environment should also be ready to begin operation as a full-time application data store.

3.4.9. Data Migration Implementation Artifacts

- Developed and fully tested
 - *Data Migration Procedures*
 - *Data Validation Procedures*
 - *Data Cleansing Procedures*
 - *Data Conversion/Transformation Procedures*
- An integrated, validated and documented *Staging Area*
- A validated and documented *Target Data Store*
- A *Data Cleansing Report*
- A *Data Conversion Report*
- *Trial Migration Results*
- *Full Data Migration Results*
- A *Parallel Operations Report*

3.4.9.1. Implementation Checklist

The MMT may use the checklist in Appendix C to ensure that all aspects of the implementation were followed during the migration. The EDS Team may use the same checklist, with supporting commentary, to evaluate the data migration.

3.4.9.2. NARA (National Archives and Records Administration) Data Validation Requirements

Based on the disposition instruction within the NARA General Records Schedule, data validation artifacts (Test/Acceptance Plan) should be kept for 5 years AFTER a system is superseded by a new iteration, or is terminated, defunded, or no longer needed for agency/IT administrative purposes.

3.5. Data Migration Closeout

3.5.1. Closeout Overview

Once the operational scope of the data migration project is finished, the final data migration documentation must be prepared and submitted to the business area stakeholders and the EDS Team.

3.5.2. Data Migration Closeout Tasks & Subtasks

The following table presents the four major tasks of the Data Migration Closeout Phase. The Data Migration Project Manager, in collaboration with the DMT, will determine the order in which these tasks will be performed.

DOCUMENT DATA MIGRATION RESULTS	DOCUMENT LESSONS LEARNED ³⁰	PERFORM KNOWLEDGE TRANSFER	COMMUNICATE DATA MIGRATION RESULTS
Document Data Migration Results	Document Lessons Learned	Perform Knowledge Transfer	Disseminate Data Migration Results
Critical Success Factors for Data Migration Closeout		Provide Stakeholder Training	Obtain Approval for Project Completion

Table 3-9: Data Migration Closeout

3.5.3. Document Data Migration Results

The MMT, in tandem with the TMT, shall compile statistics from the data migration.³¹

3.5.3.1. Critical Success Factors for Data Migration Closeout

Best Practices dictate that the following critical success factor goals be achieved while summarizing data migration results:

- Understand Data Requirements
- Ensure that all user expectations are addressed and/or satisfied by the reported results
- Establish and/or follow a standard set of report formats for disseminating migration results
- Follow standardized and/or mandated project closeout procedures
- Obtain final stakeholder approval

³⁰ *IPM Data Management Plan*

³¹ *2006 Best Practices for Data Migration*, page 8

3.5.4. Document Data Migration Lessons Learned

The MMT, in tandem with the TMT, shall document lessons learned from the data migration.³²

3.5.5. Perform Knowledge Transfer

All plans, architectures, and results should be packaged and provided to the business area stakeholders as an audit trail of the data, and as a demonstrated example of a successful migration.

3.5.6. Perform Stakeholder Training

Any stakeholder training in data migration procedures and/or the staging and target data stores may occur at this point. This training should be based on a project training plan and, is critical to ensure a smooth transition of the newly developed system to Federal Student Aid.

3.5.7. Disseminate Data Migration Results

As with the plans and designs, it is crucial to communicate the results of the migration. The results and lessons learned shall be provided to the EDS Team to evaluate the methodology employed in the migration, and to determine whether revisions to the methodology are necessary to improve future data migration efforts. The results shall also be provided to the business area stakeholders, who shall determine whether the business requirements of the migration have been met. Once the data migration results have been reviewed and approved by the business area stakeholders and the EDS Team, all documentation and artifacts should be made available to all other Federal Student Aid projects for analysis and leveraging of plans, procedures, and designs.

3.5.8. Obtain Approval for Project Completion

The approval of the data migration results by the business area stakeholders and the EDS Team is also considered the approval for project completion. The Data Migration Project Manager and the EDS Team close out the project.

3.5.9. Data Migration Closeout Artifacts

- A detailed and interpreted *Statistics Report* on Full Data Migration
- A *Lessons Learned Report*
- A *Training Plan* for Staging and Target Environment Configuration
- A *complete set of Data Migration Documents and Artifacts* in electronic and/or paper format.

3.5.10. Data Migration Closeout Checklist

The MMT may use the checklist in Appendix C to ensure that all aspects of data migration closeout set forth in this methodology are accounted for in the *Data Migration Plan* and in any data migration summary reports. Once results of the migration are submitted to the EDS Team, the same checklist may be used, with supporting commentary for feedback, to evaluate the thoroughness of the reported results.

³² 2006 *Best Practices for Data Migration*, page 8

Appendix A - Acronyms and Abbreviations

ACRONYM	APPLICABLE TERM
CDM	Conceptual Data Model
ECDM	Enterprise Conceptual Data Model
ED	Department of Education
EDD	Enterprise Data Dictionary
EDS	Enterprise Data Services
EDMMG	Enterprise Data Management Master Glossary
ELDM	Enterprise Logical Data Model
FEA	Federal Enterprise Architecture
FEAF	Federal Enterprise Architecture Framework
FIPS	Federal Information Processing Standards
IT	Information Technology
ITSS	Information Technology System Services
LDM	Logical Data Model
OS	Operational System
PDM	Physical Data Model
PESC	Postsecondary Electronic Standards Council

Table A-1: Acronyms and Abbreviations

Appendix B - Glossary

TERM	DEFINITION
Best Practice ³³ :	A management idea asserting a technique, method, process, activity, incentive or reward as more effective at delivering a particular outcome than any other technique, method, process, etc.
Column	A set of data values of the same type collected and stored in the rows of a table.
Database	A set of table spaces and index spaces.
Data Conversion ³⁴ :	The [transition] of one form of computer data to another.
Data Element	A generic term for an entity/class, table, attribute, or column in a conceptual, logical, and physical data model.
Data Migration ³⁵ :	The transferring of data between storage types, formats, or computer systems.
Enterprise Conceptual Data Model (ECDM)	One of the initial components of Enterprise Data Architecture. The first enterprise level data model developed. The ECDM identifies groupings of data important to Lines of Business, Conceptual Entities, and defines their general relationships. The ECDM provides a picture of the data the enterprise needs to conduct its business. (Reference: <i>U.S. Department of Education Enterprise Data Architecture – Enterprise Data Standards and Roadmaps.</i>)
Enterprise Data Dictionary (EDD)	One of the initial components of Enterprise Data Architecture. The EDD lists metadata objects and a complete description of the object at a sufficient level of detail to ensure that they are discrete and clearly understood. Such descriptions shall include, at a minimum, labels (names, titles, etc.) and definitions (or text descriptions), but may include additional descriptive metadata such as object type, classifications, content data type, rules (business, validation, etc.), valid and default values, etc. The EDD is the definitive source for the meaning of metadata objects. (Reference: <i>FSA-EDM</i>)
Enterprise Logical Data Model (ELDM)	A component of a maturing Enterprise Data Architecture. The second enterprise level data model developed. It is the result of merging application level data model information into the existing Enterprise Conceptual Data Model (ECDM). The ELDM extends the ECDM level of detail. (Reference: <i>U.S. Department of Education Enterprise Data Architecture – Enterprise Data Standards and Roadmaps</i>)
Enterprise Data Standards and Roadmaps (EDSG)	A component of a maturing Enterprise Data Architecture; rules and recommendations for the creation and updating of metadata objects and structures as well as for creating conceptual and physical models and schemas at both the enterprise and application level. (Reference: <i>FSA-EDM</i>)
Management Idea ³⁶ :	(a.k.a “Management fad”) A change in philosophy or operations that sweeps through businesses and institutions, and then disappears when enthusiasm for it wanes.
Operational Data Store (ODS):	An operational data store (ODS) is a type of database often used as an interim area for a data warehouse. Unlike a data warehouse, which contains static data, the contents of the ODS are updated through the course of business operations. An ODS is designed to quickly perform relatively simple queries on small amounts of data (such as finding the status of a customer order), rather than the complex queries on large amounts of data typical of the data warehouse.

³³ Derived from Wikipedia “Best Practice” (http://en.wikipedia.org/wiki/Best_practice)

³⁴ Derived from Wikipedia “Data Conversion” (http://en.wikipedia.org/wiki/Data_conversion)

³⁵ Derived from Wikipedia “Data Migration” (http://en.wikipedia.org/wiki/Data_Migration)

³⁶ Derived from Wikipedia “Management fad” (http://en.wikipedia.org/wiki/Management_fad)

TERM	DEFINITION
Schema (Data):	Any diagram or textual description of a structure for representing data. (Reference: <i>FSA-EDM</i>)
Table:	A set of related columns and rows in a relational database.
Table Space:	A portion of a database reserved for where a table will go. Table structure is the mapping of tables into table spaces.
Target Data Store	The data store where the migrated and/or transformed data will be moved.
Uniform Resource Identifier (URI)	The addressing technology for identifying resources on the Internet or a private intranet.

Table B-1: Glossary

Appendix C - Data Migration Project Review Checklist

The following checklist is used for identifying and tracking the phases and activities in completing a Data Migration Project.

DATA MIGRATION REVIEW CHECKLIST	STATUS			
	P	C	A	N/A
1. Overall				
1.1. Objective Described (data condition to be resolved)				
1.2. Business Value of Data Explained				
1.3. Included Artifacts Clear and Accurate				
1.4. Benefits				
1.4.1. Overall				
1.4.2. Benefits of hardware choices				
1.4.3. Benefits of software choices				
1.4.4. Benefits of Data Cleansing				
1.4.5. Benefits of Data Conversion				
1.4.6. Other				
2. Planning				
2.1. Project Planning				
2.1.1. Initiate Data Migration Project				
2.1.2. Establish Scope				
2.1.3. Establish Data Migration Roles & Responsibilities				
2.1.4. Develop Change Management Policies & Procedures				
2.1.5. Identify Critical Success Factors for Planning Phase				
2.1.6. Identify Risks / Constraints / Dependencies / Assumptions				
2.1.6.1. Funding				
2.1.6.2. Expertise Availability				
2.1.6.3. Personnel Availability				
2.1.6.4. Allowable Downtime				
2.1.6.5. Scheduling Constraints (Time)				
2.1.6.6. Legacy Environment Complexity				
2.1.6.7. Allowable Downtime				
2.1.6.8. Hardware Availability				
2.1.6.9. Physical Storage Incompatibility				
2.1.6.10. Application Performance Remediation				
2.1.6.11. Tolerance and Remediation of Planned Resources				
2.1.6.12. Tolerance and Remediation of Missed Requirements				
2.1.6.13. Other				
2.1.6.14.				
2.1.7. Develop Risk Mitigation Strategy				
2.2. Migration Requirements				
2.2.1. Determine Design Requirements				
2.2.1.1. Required Design Artifacts				
2.2.1.2. Migration Execution Performance Metrics				
2.2.1.3. Migration Execution Requirements				
2.2.1.4. Other				
2.2.2. Determine Time Requirements				

DATA MIGRATION REVIEW CHECKLIST	STATUS			
	P	C	A	N/A
2.2.3. Determine Technology				
2.2.3.1. Data Storage Distribution				
2.2.3.2. Physical re-location Requirements				
2.2.3.3. Target Hardware Configuration				
2.2.3.4. Target Software Configuration				
2.2.3.5. Homogeneous vs. Heterogeneous Storage				
2.2.3.6. Multi-vendor Storage Environment				
2.2.3.7. Target Data Capacity				
2.2.3.8. Other				
2.2.4. Determine Stakeholder Requirements				
2.2.5. Determine User Expectations				
2.2.6. Determine Data Security Requirements				
2.2.6.1. Source Data Protection (Recoverability)				
2.2.6.2. Access to Migrating Data				
2.2.6.3. Access to Migration Environment				
2.2.6.4. Access to Documentation				
2.2.6.5. Access to Legacy Environment				
2.2.6.6. Access to Interim Environment				
2.2.6.7. Access to Target Environment				
2.2.6.8. Other				
2.2.7. Consider Future Requirements				
1.2.1.1. Future Capacity Growth				
1.2.1.2. Durability of Target Data Storage				
1.2.1.3. Migratability of Target Data Storage				
1.2.1.4. Re-usability of Target Data Storage				
1.2.1.5. Scalability of Target Data Storage				
1.2.1.6. Other				
2.3 Current Environment				
2.3.1 Existing Data Related Artifacts				
2.3.1.1 Availability of Data Architecture Documents				
2.3.2 Blueprint Current Data Architecture				
2.3.2.1 Availability of Current Data Storage				
2.3.3 Profile Legacy Data				
2.3.4 Determine IT Infrastructure Requirements				
2.4 Develop Data Migration Plan				
2.4.1 Identify Technology Options				
2.4.2 Determine Data Migration Method				
2.4.3 Plan Data Content Management Strategy				
2.4.3.1 Data Conversion Plan				
2.4.3.2 Data Quality Metrics				
2.4.3.3 Data Loss Tolerance				
2.4.3.4 Data Quality Remediation Plan				
2.4.3.5 Data Integration / Reconciliation Plan				
2.4.3.6 Data Archival Strategy				
2.4.3.7 Other				
2.4.4 Plan Parallel Operation				
2.4.5 Plan Data Security Strategy				
2.4.6 Develop Data Migration Plan				

DATA MIGRATION REVIEW CHECKLIST	STATUS			
	P	C	A	N/A
3. Analysis & Design				
3.1 Analysis				
3.1.1 Analyze Current Environment				
3.1.2 Evaluate Data Migration Technology				
3.1.3 Evaluate Data Quality (Data Profiling)				
3.1.4 Correlate Data to Business Processes				
2.1.5 Identify Critical Success Factors for Analysis & Design Phase				
3.2 Determine Security Controls				
3.2.1 Design Enterprise Management Controls				
3.2.2 Design Operational Security Controls				
3.2.2.1 Access to Migration Environment				
3.2.2.2 Access to Documentation				
3.2.3 Design Technical Security Controls				
3.2.3.1 Access to Migrating Data				
3.2.3.2 Access to Legacy Environment				
3.2.3.3 Access to Interim Environment				
3.2.3.4 Access to Target Environment				
3.3 Design Data Environment				
3.3.1 Design Security Data Architecture				
3.3.2 Design Staging Area				
3.3.3 Design Target Data Architecture				
3.3.4 Correlate Migration Data (Source / Stage / Target)				
3.3.4.1 Legacy to Staging				
3.3.4.2 Legacy to Target				
3.3.4.3 Staging to Target				
3.3.5 Correlate Migration Data to Procedures				
3.3.6 Determine Technology Configuration (ETL, RDBMS, etc.)				
3.3.7 Design Migration IT Infrastructure				
3.4 Design Migration Procedures				
3.4.1 Design Data Security Procedures				
3.4.2 Design Data Staging Procedures				
3.4.3 Design Data Cleansing Procedures				
1.4.4 Design Data Conversion Procedures				
3.4.5 Design Target Data Migration Procedures				
3.4.6 Design Data Validation Procedures				
3.4.7 Design Data Quality Remediation Procedures				
3.4.8 Refine Data Migration Test Plan				
3.4.9 Design Data Quality Reports and Process for Reconciliation				
Implementation				
3.5 Develop Migration Procedures				
3.5.1 Configure Resources				
3.5.1.1 Tool Configuration				
3.5.1.2 Application Configuration				
3.5.1.3 Data Configuration				
3.5.2 Develop and Test Data Migration Procedures				
3.5.3 Develop and Test Data Quality Validation Procedures				
3.5.4 Develop and Test Data Cleansing Procedures				
3.5.5 Develop and Test Data Conversion / Transformation				

DATA MIGRATION REVIEW CHECKLIST	STATUS			
	P	C	A	N/A
Procedures				
3.5.6 Establish Access to Staging and Target Area				
3.5.7 Identify Critical Success Factors for Implementation Phase				
3.6 Stage Data				
3.6.1 Create Staging Area				
3.6.2 Populate Staging Area				
3.6.3 Integrate Staged Data				
3.6.4 Validate Staged Data				
3.7 Cleanse Data				
3.7.1 Cleanse Data				
3.7.2 Validate Cleansed Data				
3.8 Convert / Transform Data				
3.8.1 Convert / Transform Data				
3.8.2 Validate Converted / Transformed Data				
3.9 Migrate Data				
3.9.1 Perform Trial Migration				
3.9.2 Validate Results of Trial Migration				
3.9.3 Obtain Approval for Full Migration				
3.9.4 Perform Full Migration				
3.9.5 Validate Full Migration				
3.10 Post-Migration				
3.10.1 Operate legacy and target environment in parallel				
3.10.2 Validate parallel operation and data				
3.10.3 Release Data Environments				
4. Data Migration Closeout				
4.1 Document Data Migration Results				
4.2 Identify Critical Success Factors for Data Migration Closeout Phase				
4.3 Document Data Migration Lessons Learned				
4.4 Perform Knowledge Transfer				
4.5 Communicate Data Migration and Lessons Learned				

Table C-1: Data Migration Review Checklist

Legend: P – Planned
 C – Completed
 A – Accepted
 N/A- Not applicable

Appendix D - Data Migration Project Deliverables

DATA MIGRATION DELIVERABLES		STATUS			
DELIVERABLE	DESCRIPTION	P	C	A	N/A
1. Planning					
1.1. Project Management Plan					
1.1.1. R / C / A / D	Documentation outlining the Risks, Constraints, Assumptions, & Dependencies associated with proposed data migration.				
1.1.2. Risk Mitigation Plan	Documentation describing the plan for preventing or responding to the risks associated with the data migration; including a matrix demonstrating the likelihood & impact of occurrence of specific risks.				
1.1.3. Project WBS	Documentation showing planned project activities and the time and personnel resources to be applied to performing those activities.				
1.2. Data Migration Plan	Plan outlining how the data migration from the data source to the target system is planned for. It includes the plan for ensuring that post-migration data content satisfies the requirements of the target data environment. This document supplements the overall Project Plan and consists of multiple deliverables:				
1.2.1. Data Conversion Plan	The plan for converting the form and content of source data into satisfactory target data.				
1.2.2. Data Quality Metrics	Documentation describing the standards of measurement to be applied to the content data before, during, and after migration. This document includes information regarding the acceptable level(s), if any, of lost data content or meaning as a result of source data undergoing a change in content or form.				
1.2.3. Data Quality Remediation Plan	The plan for correcting any identified data quality issues impacting or resulting from the data migration. <i>(RE: 3.4.4.4.3 Data Migration Plan)</i>				
1.2.4. Integration / Reconciliation	The plan for integrating and reconciling all the different source data system into one set of satisfactory target data.				
1.2.5. Data Migration Test Plan	The plan for validating the successful movement of source data to the new target data store.				
1.2.6. Data Archival Strategy	The strategy for archiving historical data that is no longer needed.				
1.3. Change Management Plan	Documentation showing the process by which version control of project documentation and configuration management will be performed such that all results meet the highest reasonable expectations of quality.				
1.4. QA Plan	Documentation showing the process by which the project team shall ensure that all activities are performed such that all results meet the highest reasonable expectations of quality.				
1.5. Communications	Documentation showing how information shall be				

DATA MIGRATION DELIVERABLES		STATUS			
DELIVERABLE	DESCRIPTION	P	C	A	N/A
Plan	communicated among the members of the project team and between members of the project team and external stakeholders.				
1.6. Data Security Plan	Documentation showing how the data security will be implemented and adhered to. This document could reference existing Federal Student Aid Data Security policies and procedures.				
1.7. Data Migration Requirements	Documentation describing conditions that must be met in order to deem the data migration successful.				
1.8. Parallel Operation Plan	The plan for validating the successful movement of source data to the new target data store through parallel operation of the old and the new system over a defined period of time.				
2. Analysis & Design					
2.1. Data Profile Report	Documentation showing the data quality challenges assessed through the <i>Data Profiling Report</i> .				
2.2. Data Quality Report (Pre-migration)	Documentation showing how the data quality is assessed and the planned remediation.				
2.3. Security Controls	Documentation showing how the data security will be implemented and adhered to.				
2.4. Migration Data Architecture	Documentation showing how the data architecture is designed at the logical and physical level through ERDs.				
2.4.1. Staging Area LDM	Staging Area Logical Data Model: Graphical representation of the information and business rules of the staging area. (Note: <i>representation should comply with FSA standard modeling methodology.</i>)				
2.4.2. Staging Area PDM	Staging Physical Data Model: Graphical representation of the internal data structures and constraints of the staging area. (Note: <i>representation should comply with FSA standard modeling methodology.</i>)				
2.4.3. Staging Area DD	Staging Data Dictionary				
2.4.4. Target LDM	Target Logical Data Model: Graphical representation of the information and business rules of the target data store. (Note: <i>representation should comply with FSA standard modeling methodology.</i>)				
2.4.5. Target PDM	Target Physical Data Model: Graphical representation of the internal data structures and constraints of the target data store. (Note: <i>representation should comply with FSA standard modeling methodology.</i>)				
2.4.6. Target DD	Target Data Dictionary				
2.4.7. Data Migration Activity Model	Graphical representation of the activities / functions to be performed during the data migration. (Note: <i>representation should comply with FSA standard modeling methodology.</i>)				
2.4.8. Data Migration	Documentation outlining who has access in what capacity to which environment.				

DATA MIGRATION DELIVERABLES		STATUS			
DELIVERABLE	DESCRIPTION	P	C	A	N/A
CRUD Matrix					
3. Implementation					
3.1. Fully developed and tested Data Migration Procedures	Documentation and code of the data migration procedures and related test results.				
3.2. Fully developed and tested Data Validation Procedures	Documentation and code of the data validation procedures and related test results.				
3.3. Fully developed and tested Data Cleansing Procedures	Documentation and code of the data cleansing procedures and related test results.				
3.4. Fully developed and tested Data Conversion/Transformation Procedures	Documentation and code of the data conversion/transformation procedures and related test results.				
3.5. Data Cleansing Report	Documentation of data cleansing findings.				
3.6. Data Conversion Report	Documentation of the executed data conversion.				
3.7. Trial Migration Results	Documentation of the executed trial data migration(s).				
3.8. Acceptance / Approval Documentation	Documentation demonstrating the stakeholder approval of the data migration procedures readiness for deployment to production.				
3.9. Full Migration Results	Documentation of the results of the full data migration.				
3.10. Parallel Operations Report	Documentation outlining the results of the parallel operation of the old and new data store. This report enables stakeholders to decide whether the parallel operation can be ended.				
4. Close-out					
4.1. Data Migration Results	<p>Documentation describing</p> <ul style="list-style-type: none"> • Statistics of the data migration such as actual data quality and volume measurements, downtime, data loss, etc.; • Unresolved issues; • <i>[Others]</i> <p>The actual content of this artifact shall be determined by the overall scope of the data migration. Complex migration efforts would naturally require more extensive reporting, while more basic migrations may require less content.</p> <p>Please note: Based on the disposition instruction within the NARA General Records Schedule, data validation artifacts (Test/Acceptance Plan) should be kept for 5 years AFTER a system is superseded by a new iteration, or is terminated, defunded, or no longer needed for agency/IT administrative purposes.</p>				

DATA MIGRATION DELIVERABLES		STATUS			
DELIVERABLE	DESCRIPTION	P	C	A	N/A
Data Migration Lessons Learned	Documentation describing lessons learned during the project that may explain results and contribute to enhancements to future data migration efforts.				
Training Plan	Documentation on the knowledge transfer between the data migration team and the stakeholders take place, including the relevant training material.				
Complete set of all project deliverables	The compiled inventory of all artifacts resulting from the data migration project.				

Table D-1: Data Migration Deliverables

Legend: **P** – Planned
C – Completed
A – Accepted
N/A- Not applicable