



CDC UNIFIED PROCESS PRACTICES GUIDE



DATA CONVERSION

Purpose

The purpose of this document is to provide guidance on the practice of **Data Conversion** and to describe the practice overview, requirements, best practices, activities, and key terms related to these requirements. In addition, templates relevant to this practice are provided at the end of this guide.

Practice Overview

At the highest of levels data conversion/migration is defined as the process of translating data from one format to another. It involves the planning of steps and mapping of data fields to convert one set/type of data into a different, more desired, format. Data conversions and/or migrations may be performed for a number of reasons such as hardware or software upgrades, business expansions, data center moves, combining multiple systems into one, normalization of data, etc. Planning for a data conversion/migration also requires reviewing existing business processes, organizational policies and procedures, security, etc., identifying areas that may be impacted by variances between the old and new systems, and planning accordingly to deal with such impacts.

Often the successful implementation of a new system is dependant upon the ability to convert data from the old system to the new system. An example of data conversion/migration may be as simple as converting WordPerfect documents into Microsoft Word documents or as complex as migrating entire databases from one application, and schema, to another.

Converting and migrating data often involves both software and human intervention. This is especially true when migrating data contained within a database. Whether a project team uses commercial-off-the-shelf software to assist the effort, develop their own in-house solution, or uses humans to do the work manually is dependant upon the systems and data types being converted. Regardless of how the effort is performed it's important to recognize that information can easily be discarded but is difficult, if not impossible, to restore. It is especially important to evaluate this, and understand how data will be impacted, when converting from one data type to another. The "Practice Activities" section of this document outlines a high-level strategic approach that may be used to prepare for a data conversion/migration regardless of the approach taken.

Best Practices

The following best practices are recommended for **Data Conversion**:

- **Risk** – Identify potential project risks related to data conversion as early in the project life cycle as possible. Document these initially identified risks in the project charter and clearly communicate their potential consequences to project sponsors and stakeholders.
- **Plan** – Careful planning is the most important contributing factor to a successful project.
- **Communicate** – Prevent confusion and/or misinterpretation by effectively communicating every detail, and step, of the planned migration/conversion process.
- **Stakeholders** – Involve business owners and other relevant stakeholders.
- **Backup** – System backups should be taken incrementally to allow the project team to revert back to any point in the system's life.
- **Test Sample** – Perform test conversions/migrations with a sampling of the data before attempting to apply the solution to the entire system.
- **Full Test** – Perform a controlled, full-volume, "dress rehearsal" test, of the activities required when converting data to the target system. This is an end-to-end test of the entire data conversion process and the data on the new system. It includes testing the processes and procedures planned for the conversion, the new system data, business rules, technology, etc. This may be less important in small low-risk project but is especially important in large high-risk projects where many users may be impacted by unforeseen events resulting from the migration.
- **Validate** – Validate/reconcile that the converted data is accurate and complete.



CDC UNIFIED PROCESS PRACTICES GUIDE



DATA CONVERSION

Practice Activities

For software development projects the following practice activities are appropriate:

1. Planning for Data Conversions

- The most important factor to successful data conversion/migration is careful planning and effective communication of every detail, and step, of the process.
- Inventory the system(s) and supporting components to understand what it is that the project team will be working with as they transition the data.
- Review the data and data types being converted. Its important to understand information such as:
 - The amount, type, and quality of data
 - The original and target sources and formats
 - Any cross-reference complexities
- Evaluate the experience of the project team to determine their ability to successfully perform the data migration/conversion. It may be necessary to hire additional resources or outsource some, if not all, of the work if the appropriate skill set is not available in-house.
- Identify the criticality of the data. This may impact the approach taken to convert/migrate the data as well as the amount and type of resources required to successfully perform the effort.
- Determine if the most appropriate, low risk, approach is to perform the migration in-house or to outsource the effort or a combination of the two. Each option has its own advantages and disadvantages. Some advantages of performing this effort in-house may include control and security of data, schedule and resource flexibility, and possible cost savings. Outsourcing such efforts will cost money but often brings a level of expertise not always available in-house. Outsourcing also allows for internal resources to be allocated towards other efforts and priorities.
- Determine how the data conversion/migration will be performed. Is there a requirement to run parallel system, will their be a one time cut-over to the new system, archive the old system or keep it running, etc
- Analyze the above information as inputs into the conversion/migration process to help determine costs, schedules, software needs, and any required human intervention.
- Perform a high-level mapping to determine which data elements in the existing system will be converted/migrated to the new system. Decide which data will be transferred, converted, which is redundant, etc.
- Develop business rules that outline how items will be handled. Items such as blank records, new codes, inappropriate entries, etc.
- Develop conversion scripts, as needed. Conversion scripts are used for extracting data from the source, transforming the data as needed, and loading the data into the target.
- Choose the best human and/or software approach to maximize quality and minimize expense.
- Develop a schedule that maps out exactly how the conversion/migration is expected to happen.
- Create a specification document that maps out exactly how the converted data will look.
- Other planning considerations should include items such as communication, education, data normalization, quality assurance, and validation of data accuracy and completeness.

2. Performing Data Conversions

- Generate a backup of all data prior to any manipulation or migration. This backup represents the system baseline prior to any human and/or software interaction with the system or system data outside of the normal operating processes. If needed, this backup can be used to restore the system. System backups should be taken incrementally while stepping through the process of preparing, moving, and manipulating data. This is done to allow the project team to revert back to any point throughout the process that they identify as correct if for some reason they run into issues during later steps.
- Extract test data from the legacy system.
- Normalize the test data. Often one of the main goals of performing a data conversion/migration is to combine multiple data sources into one standardized format. This is referred to as normalizing data. Data is often normalized by structuring database tables logically so that they contain



CDC UNIFIED PROCESS PRACTICES GUIDE



DATA CONVERSION

information related only to the items within that table and then linking/joining tables appropriately in order to build the functionality desired by the database user. This is done to minimize unnecessary redundancy and increase data efficiencies.

- Perform a test conversion of a sample of existing data and make adjustments if necessary.
- Depending on the criticality of the system, one or more mock conversions may also be necessary. A mock conversion is a controlled “dress rehearsal” of the execution activities required when converting data into the target system. It is meant to be a pre-go-live test in that everything that occurs in a go-live conversion has been tested in a mock conversion. The main objective of the mock conversions is to test the conversion process and scripts. The mock conversions are intended to identify and resolve any conversion software issues, address any configuration issues, identify any additional data validation and verification efforts, and prove the conversion procedure. Each mock conversion will simulate the real go-live process with actual data volumes.
- Once the project team is confident that the data conversion should go smoothly plan and communicate a date appropriate for the full conversion to take place. It is often most appropriate to perform this during non-business hours, preferably on a Friday evening. This allows a few days of contingency (over the weekend) to resolve any last minute issues or to revert back to the system backup taken just before the conversion/migration began.
- Normalize the system data.
- Initiate the data conversion.

3. Validating/Evaluating Data Conversions

- Validate/reconcile the converted data for accuracy and completeness. Check items such as:
 - Formatting of data elements
 - Data completeness
 - Data accuracy
- Eliminate duplicate elements.
- Resolve any issues that may.

Practice Attributes

This section provides a list of practice attributes to help project teams determine when and how development of a **Data Conversion** impacts a project.

Practice Owner	CDC UP Project Office
Criteria	All projects that require the conversion or migration of data, regardless of type or size, should have some type of document outlining the steps and processes that will be performed in order to successfully move the data.
Estimated Level of Effort	Significant
Prerequisites	N/A
Practice Dependencies	N/A
Practice Timing in Project Life Cycle	N/A
Templates/Tools	N/A
Additional Information	N/A

Key Terms

Follow the link below to for definitions of project management terms and acronyms used in this document.
<http://www2.cdc.gov/cdcup/library/other/help.htm>



CDC UNIFIED PROCESS PRACTICES GUIDE



DATA CONVERSION

Related Templates/Tools

Below is a list of template(s) related to this practice. Follow the link below to download the document(s).
<http://www2.cdc.gov/cdcup/library/matrix/default.htm>

- N/A