

Data Conversion Plan

*<Project or System Name>*

**U.S. Department of Housing and Urban Development**

*<Month, Year>*

# Document History

<Provide information on how the development and distribution of the Data Conversion Plan is controlled and tracked. Use the table below to provide the version number, date, author, and a brief description of the reason for creating the revised version.>

|  |  |  |  |
| --- | --- | --- | --- |
| **Version No.** | **Date** | **Author** | **Revision Description** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Contents**

[Document History 1](#_Toc378435272)

[1. Assumptions/Constraints/Risks 3](#_Toc378435273)

[1.1 Assumptions 3](#_Toc378435274)

[1.2 Constraints 3](#_Toc378435275)

[1.3 Risks 3](#_Toc378435276)

[2. Data Conversion Strategy 4](#_Toc378435277)

[2.1 Conversion Scope 4](#_Toc378435278)

[2.2 Conversion Approach 4](#_Toc378435279)

[2.3 Roles and Responsibilities 5](#_Toc378435280)

[2.4 Conversion Schedule 5](#_Toc378435281)

[2.5 Data Quality Control Strategy 6](#_Toc378435282)

[3. Data Conversion Preparation 7](#_Toc378435283)

[3.1 Prerequisites 7](#_Toc378435284)

[3.2 Backup Strategy 7](#_Toc378435285)

[3.3 Restore Process 7](#_Toc378435286)

[4. Data Conversion Specifications 8](#_Toc378435287)

[Appendix A: References 9](#_Toc378435288)

[Appendix B: Key Terms 10](#_Toc378435289)

# Assumptions/Constraints/Risks

##  Assumptions

<Describe any assumptions or dependencies regarding the data conversion effort. These may concern such areas as related software or hardware, operating systems, end-user characteristics, and/or the data that must be available for the conversion.>

## Constraints

<Describe any limitations or constraints that have a significant impact on the data conversion effort. You may have constraints imposed by any of the following (the list is not exhaustive):

* *Hardware or software environment*
* *End-user environment (user work and delivery schedules, timeframes for reports, etc.)*
* *Availability of resources*
* *Interoperability requirements (e.g. the order that data is processed by each system involved in the conversion)*
* *Interface/protocol requirements*
* *Data repository and distribution requirements (e.g. volume considerations, such as the size of the database and amount of data to be converted; the number of reads and the time required for conversions)*
* *Referential data integrity*
* *Time allowed to complete the conversion process>*

## Risks

<Describe any risks associated with the data conversion and proposed mitigation strategies. Include any risks that could affect conversion feasibility, technical performance of the converted solution, the conversion schedule, costs, backup and recovery procedures, etc.>

# Data Conversion Strategy

## Conversion Scope

<Provide a rationale for the conversion and a general description of the boundaries of the data conversion effort. This may include, but not be limited to, specific system functions affected and functions/data not affected/converted. Provide a high level mapping of the data and data types to be converted or migrated to the new system (e.g. the amount, type, and quality of the data; the original and target sources and formats; and any cross-reference complexities).>

## Conversion Approach

<Describe the approach used to extract, transform, cleanse, and load data from the source to target destinations during the conversion/migration process. Consider and address the following in this section and/or appropriate subsections, if applicable:

* *Identify if the conversion process is implemented in phases or stages, and if so, identify which components will undergo conversion in each phase*
* *Identify what data related to specific business processes must be converted first*
* *Describe any automated data conversion tools used (e.g. extract, transform, and load (ETL) tools)*
* *Identify and describe any part of the conversion process performed manually*
* *Identify and describe any custom-developed conversion programs that are needed, and associated performance tuning*
* *Identify testing requirements, test plan and criteria for a Go/No-Go decision*
* *Identify staffing approach*
* *Identify if parallel runs of the old and new systems will be necessary during the conversion process, or if there will be a one-time cutover to the new system*
* *Identify whether data availability and use should be limited during the conversion process*
* *Describe security and privacy controls required for the conversion process*
* *Describe the disposition of obsolete or unused data that is not converted*
* *Identify the retention policy for the data that is converted in case of fallback and a need to rerun the conversion process*
* *Use the HUD Data Dictionary column “Data Source System” to identify the data to be migrated and to where. Provide a full inventory. If the data element is not to be migrated then mark the data element as “not migrated”*
* *Consider National Archives and Records Administration (NARA) and eDiscovery retention policies>*

## Roles and Responsibilities

<List all stakeholders and document their roles and responsibilities in the conversion process.>

| Role | Name | Responsibilities | Contact Info |
| --- | --- | --- | --- |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Table 2.1 Roles and Responsibilities**

## Conversion Schedule

<Provide a schedule of conversion activities accomplished in accordance with this Data Conversion Plan. Show the required tasks in chronological order, with beginning and ending dates of each task, the key person(s) responsible for the task, dependencies, and milestones.

Some key considerations to consider in the schedule include:

* *Since data conversion often relies on the delivery of files from outside sources, it is imperative that the project team requests these files as soon as they have identified the data sources. Some stakeholders will have to implement a Memorandum of Understanding (MOU), before data can be provided for the data conversion*
* *The expectation of when the data conversion tests and files will be ready to load into an independent verification and validation (IV&V), user acceptance testing (UAT) or test center environment to support end-to-end testing*
* *The identification of when the production data conversion files will be provided for production deployment - the timing of these files and the length of time required to load the files is a critical consideration*
* *Identification of the type of control measures you will use to ensure that data from other systems are properly loaded into new solution*
* *Parallel testing impacts to the data conversion schedule*

If appropriate, use tables and/or graphics to present the schedule. Ensure that the information is appropriately integrated into the overall project schedule. The schedule should be as comprehensive as possible; however, the schedule may be revised as needed at later points in the life cycle. Rather than providing this schedule in the table below, the schedule may be added as an appendix and may be developed in a project management tool.

This schedule may be incorporated into the Project Schedule (WBS) created in the Planning Phase.>

| **Task #** | **Task Description** | **Begin Date** | **End Date** | **Key Person(s) Responsible** | **Dependencies** | **Milestone** |
| --- | --- | --- | --- | --- | --- | --- |
| <task #> | <task description> | <mm/dd/yy> | <mm/dd/yy> | <name(s)> | <task #(s)> | <Yes/No> |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

Table 2.2 - Data Conversion Schedule

## Data Quality Control Strategy

<Describe the strategy used to ensure data quality before and after all data conversions. Also, describe the approach to data scrubbing and quality assessment of data before they are moved to the new or converted system. Describe the manual and/or automated controls and methods to be used to validate the conversion and to ensure that all data intended for conversion have been converted. Describe the process for data error detection and correction, and the process for resolving anomalies.

Identify the types of data quality problems that may occur, including, but not limited to the following considerations:

* *data type redefinitions (e.g. alphas in dates and numbers, embedded information in codes and intelligent keys, implied content)*
* *garbled content (e.g. multiple uses for a single field, freeform text values, corrupted data, un-initiated data)*
* *invalid record relationships (e.g. broken chains in set relationships, orphan records (on natural key), mismatched keys (set vs. natural key))*
* *invalid content (e.g. values out of defined range, code fields not on a valid list of values or lookup table, blank fields (optionality), inconsistent use of defaults)*
* *context changes (e.g. import of external data, historic changes to operational parameters (system upgrades), synchronization timing of duplicated de-normalized data)*
* *behavior issues (e.g. variations in actual data from planned constraints of size, data type, validation rules, and relationships)>*

# Data Conversion Preparation

## Prerequisites

<Describe all preparatory and/or initiation processes that must be completed prior to data conversion. Describe specific data preparation requirements. If the data will be transported from the original system, provide a detailed description of the data handling, conversion, and loading procedures. If the data will be transported using machine-readable media, describe the characteristics of that media. Identify any support materials needed for the conversion process.>

## Backup Strategy

<Describe the process to create and manage the source and target data baselines prior to any manipulation or migration. Also describe backups that may occur incrementally while stepping through the process of preparing, moving, and manipulating the data during conversion.>

## Restore Process

<Describe the process to restore the source data if the need to revert to a previous back-up is identified at any point during the conversion process.>

# Data Conversion Specifications

<Provide a cross reference of the input (source) data that is converted to the resultant output (target) data. Also, identify if any of the data are derived from other data. Provide transformation/cleansing rules for each data element and any other additional considerations. Transformation and cleansing rules may include, but are not limited to, the following:

*Translation of literal value(s) to literal value(s)*

*Default null to literal value*

*Empty field processing (i.e. null to space or space to null)*

*Formulas (i.e. simple equations and mathematical expressions)>*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **Source Data Element** | **Destination** | **Target Data Element** | **Transformation/ Cleansing Rules** | **Notes** |
| <Source location (e.g. system/file/ database table> | <Source data element identifier (e.g. SSN)> | <Target location (e.g. database table)> | <Target data element identifier (e.g. member ID)> | <Describe data transformation that is to occur, including any data cleansing> | <Describe any timing constraints or anything unique about the conversion> |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

Table 4.1 - Data Conversion Specifications

# Appendix A: References

<Insert the name, version number, description, and physical location of any documents referenced in this document. Add rows to the table as necessary.>

Table A.1 below summarizes the documents referenced in this document.

|  |  |  |
| --- | --- | --- |
| **Document Name** | **Description** | **Location** |
| <Document name and version number> | <Document description> | <URL to where document is located> |
|  |  |  |
|  |  |  |

Table A.1 - Appendix A: References

# Appendix B: Key Terms

Table B.1 below provides definitions and explanations for terms and acronyms relevant to the content presented within this document.

|  |  |
| --- | --- |
| **Term** | **Definition** |
| <Insert Term> | <Provide definition of term and acronyms used in this document> |
|  |  |
|  |  |

Table B.1 - Appendix B: Key Terms