

Safety Data Migration

Data migration is an exercise that is often overlooked as a simple step of copying/moving data between retiring an existing system and bringing to life a new one. The actual task of data migration is a complex one requiring experience and a detailed understanding of business processes and technical implementation.

Data Migration is one of the most critical and challenging part of any drug safety system implementation or an upgrade program. Strict regulatory compliance and sensitivity towards data is required to migrate the legacy data seamlessly for future case processing.

Pharmaceuticals and Contract Research Organizations(CRO) spend millions of dollars on data migration projects as it is the most time consuming and critical activity in safety implementation lifecycle. Industry experience indicates more than 60% data migrations run overtime or over budget. This document shall try to uncover the reasons behind these overages and explore methods to reduce Cost of Quality (CoQ).

Why Data Migration?

- Process improvement
- Adoption of a new system
- Changing regulatory requirements
- Market consolidation or Product acquisitions

How it works?

Data migration starts with source system analysis and extensive data profiling to uncover data quality issues at an early phase of a project. Next step is to prepare source to target mapping capturing the functional requirements. An ETL tool or a custom program can extract the source data and apply data conversion as per the functional requirements. This thorough analysis provides an opportunity to business users to identify data exceptions and perform data fixes prior to the complete load. Finally, data validation is performed to verify the migrated data.

Challenges

Project Scoping

Generally data doesn't get the attention it deserve and scoping of data migration takes backseat. Later in the project lifecycle, data quality issues severely impacts the project schedule as well as the budget.

Data profiling

Typically data quality issues are caught late in the project lifecycle, although the expectation is to catch them at the beginning of the project. Thorough data profiling can uncover data issues at an early phase. There is a tendency of using data profiling queries from previous projects which can paint a misleading picture.

Lack of knowledge of disparate safety systems

Safety data migration projects require expertise on multiple safety systems including technical understanding along with domain knowledge. These skills are niche and scarce.

Data validation

Sampling of legacy data for validation is one of the major challenges for the quality team. A good sample must represent all data scenarios in the legacy system.

Methodology

Absence of structured methodology in data migration projects can result in stressful conditions for business and project team. The business and development team usually work in silos and communication of critical decisions become a challenge.

Volume of cases

Legacy systems carries a lot of data. This ever increasing amount of data become a challenge for performance bench marking, IT planning and hardware sizing.

Solution

Plan

Data migration projects are complex, business should aim to migrate what is required and feasible. After scope refinement, the project time line and budget should be put in place or reworked.

Profile

Data profiling should be performed to uncover data quality issues and informed decisions can be made for manual or

programmatic data cleansing. Data profiling would be helpful for business users to decide what needs to be migrated. Data profiling can reduce cost of iterations and rewriting of code.

Design

Develop mapping specification which includes all data conversion rules i.e. direct mapping, code-list mapping, custom business rule. Create a design that meets the requirements, system architecture and data volume.

Develop

Develop the code for Extract, Transform and Load (ETL) process as per mapping specification and verify each rule. Multiple test runs will rapidly identify the errors and help in making key decisions for the project.

Validate

Business user must validate data by report comparison, case processing, follow-up processing, and automated test results.

Execute

In execution, ETL program will extract the source data, convert data as mapping specifications and loads into target system.

Data quality matrices should be prepared regularly during planning, profiling, designing, development and validation phases. These matrices provides insight to business users and indicate data quality improvement over time.

Tools and Approaches

There are many standard ETL tools available in the market that can be used to perform safety data migration. However, custom program are preferred as data migration is a onetime activity and they are also cost effective in comparison to ETL tools.

Typically, there are three approaches that can be adopted based on size and the complexity of source data

Table by Table

- Loads data table by table
- Faster execution time
- Shorter development cycle
- Suitable for large migrations (>1Million cases)
- Case level error tracing not possible
- Complex translation layer development
- May require temporary target system structure changes

Case by Case

- Loads data case by case
- Case level error tracing possible
- Simpler translation layer
- No impact on target system object structure
- Suitable for source with poor data quality
- Suitable for small and medium size migrations (up to 1 Million cases)
- Comparatively slow execution time
- Long development cycle

Hybrid (E2B plus custom SQL)

- E2B files are used for migration
- E2B files standardize data specifications across the systems
- E2B covers most of regulatory data
- Around 30% of data fields are not covered in E2B and requires custom program to migrate

Summary

For success, data migrations requires the attention they deserve. Without proper planning and data profiling, there is a high risk that the project will go over budget/over time, or even completely fail. Following 3 points are key to success in any data migration project.

Profile. Find unexpected scenarios before the migration

Improve. Remediate data quality issues during or before the migration

Monitor. Prepare and publish the data quality matrices regularly to business users

Please contact us by email (sales@vitrana.com) or phone (+91 120 463 8700) to understand more about how Vitrana™ can help your organization come up with Data Migration strategy that fits your needs.