

# Data Management

“Data Migration” services from TechUnified will provide seamless migration and ROI within few months”

JVs and Branches:  
India, United Kingdom,  
Australia, Bahrain,  
Japan, Singapore

### Introduction

This paper discusses TechUnified's approach to data integration from disparate sources and data migration from one System to another. Integration is one of the most critical items in enterprise IT budgets. The compelling business case for a single, global view of customers, products and company financials is universally accepted, but the challenges to achieving this goal are formidable.

In today's rapidly evolving business and regulatory environments, organizations must connect an increasing number of legacy, packaged and custom applications into a seamless information infrastructure to:

- Support changes and evolving business processes
- Support new regulatory requirements
- Provide accountability through operation efficiency
- React more quickly to competitive and market pressure
- Reduce IT costs

### Data Migration Planning

The data migration phase of an application implementation, switching from one system to another, upgrade or instance consolidation can extend up to several years and requires rigorous project planning. Detailed project planning is required because enterprises have traditionally underestimated the challenges involved in the data migration process, including the high cost of system maintenance and administration, the cost of developing and documenting custom interface tables and addressing complex requirements including data quality. Because data migration is such a complex process, it is important to view data migration as an entirely separate project. The ideal data migration project plan is dependent on a number of factors including:

- Incorporating best of breed data integration tools and methodologies
- Taking into consideration the full application lifecycle, including upgrades and synchronization
- Understanding the requirements of the business, then setting and communicating clear project expectations
- Carefully selecting project leaders, roles and responsibilities

## Data Migration Approach

### Data Migration Challenge

Data migration is complicated because of the need to transform data from a wide variety of transactional, legacy, relational, third party, real time and historical data sources into a new schema and supporting data types. This requires the data to be:

- Profiled and extracted from current systems, applications, databases and files
- Cleansed of incorrect, redundant and outdated records
- Transformed to support new data types, definitions, taxonomic and schemas
- Tested and validated to ensure migration routines perform according to defined specifications
- Loaded into the new application environment

The vital processes are further complicated due to the large number of disparate systems that needs to be tied together to support application implementations or instance consolidation projects.

### Data Migration Solution

The first step in evaluating data migration solution is to determine the scope of the overall effort. It is useful to review the following factors:

- Number and complexity of legacy Systems
- Type of migration (i.e. batch, one –time, synchronous or real time)
- Cost estimation for the full project lifecycle (i.e. upgrades, maintenance, customization and training)
- Legacy data quality
- Amount of data and history to be converted
- Target application architecture (i.e. schema, table, data types)
- Resource bandwidth and availability
- Budget

The next step for the IT organizations is to select the most appropriate approach given the project scope. Flexibility, reusability and total lifetime costs are the key elements to consider.

Traditionally, the enterprises have been forced to choose between developing custom code and relying on an application vendor's proprietary data migration architecture. Both the approaches present severe challenges to cost-effective, timely and successful implementations.

## Data Migration with Custom Code

Custom code can initially offer some degree of flexibility. However, as the number and complexity of integration points increases, all the hidden costs make the approach unviable. The cost in this case include creation and administration of up to 1,000 interface tables, manual data quality management process and the inability to easily adapt to changes. Another major weakness of custom code approach is the lack of reuse—most, if not all, of the interfaces created during the original implementation cannot be reused.

## Data Migration with Application Vendor Solution

An application vendor's proprietary data migration solution can provide some cost benefits, but this approach, by definition, sacrifices flexibility in favor of limited functionality within closed “black box” architecture. It also covers significant risks because vendors focus on delivering new application functionality rather than supporting data migration utilities, which can lag behind in terms of availability, reliability and depth. This can impact deployment schedules and resource allocation commitments.

## Informatica Approach

Informatica offers a third approach to data migration. The informatica solution is a proven, end-to-end data integration solution capable of supporting any data migration initiatives, regardless of size and complexity. The section below gives the complete detail of this solution

## The Informatica Data Integration Platform

This section describes the capabilities of the Informatica platform:

### Open Flexible Architecture

The Informatica data integration platform is an open standard based platform that runs on UNIX, NT and LINUX and enables enterprises to leverage all IT infrastructure investments. This includes investment in software, storage, hardware, resources, security and database platforms. It supports native access into virtually every data source or target, including, XML, EAI, legacy, relational, ERP, CRM and SCM applications.

### Meta data centric Approach

By abstracting definitions, procedures and integration logic during the integration process, metadata simplifies and accelerates data migration efforts in several ways. Comprehensive metadata:

- Promotes reuse at every level
- Simplifies data validation
- Optimizes network and processor loads
- Documents all migration process
- Captures and stores integration logic

Informatica leverages metadata across the entire data migration lifecycle by making all integration routines, including source and target data definitions, integration logic. By reusing vital integration components, enterprises are able to leverage fully validated, debugged and tested data integration processes. This helps drastically reduce data migration project timelines and integration costs.

### Flexibility Enables Migrations in Batch or Real Time

The two most common approaches employed by enterprises for the data migration are:

- **Big Bang Approach:** The big bang approach takes the production environment offline for a brief cutover period so that data can be migrated into a new application.
- **Staged Approach:** New or upgraded application modules are brought online one or two at a time

Both approaches are mostly batch driven, after migration; real time functionality is often required to synchronize data between systems for a period of up to a full year. The Informatica platform is uniquely equipped to meet all batch or real time data migration requirements, offering a wide variety of features to help ensure maximum throughput. To support real time data migration and synchronization requirements, Informatica leverages its Zero Latency (ZL) engine, which provides always-on, always listening sessions.

## Built in Data Quality Feature

As the enterprises migrate data into new applications, it needs to guarantee that only accurate, consistent and complete data is loaded. Incomplete, inaccurate, missing and duplicate data field must be identified and normalized prior to loading. Informatica comes up with built in data quality features to handle the “dirty data” problems. Informatica offers a comprehensive library of data transformations and a wide variety of exception reporting and validation features.

## Iterative Methodology

Data migration projects are extremely complex, highly unpredictable and resource intensive so the complete process takes years to complete. Project specifications shift constantly due to unanticipated complexity, changing business requirements and unfamiliarity with underlying sources and destinations.

Using Informatica object based architecture, developers are able to visually map and compare source and target data definitions, quickly develop, test and validate data migration processes from source to target systems and batch migration components into logically constructed groups and subgroups.

## Built in Transformation

Transformations are a fundamental requirement for most data migration efforts. When moving from source to target applications, database systems or file formats, data must be normalized and made consistent with the target environment.

Informatica data integration platform offers a comprehensive library of rich transformation objects, which reduce or entirely eliminate the need to write custom code and perform most complex trans-formation requirements. Informatica also offers a wide variety of productivity tools to shorten development cycles, including drag-and-drop GUI interface, workflows, visual debugger, automated documentation and developer collaboration tools