



## A discussion on how to be successful at data migrations in a world where most projects fail or significantly exceed their budgets.

## Why should I care about data migration?



The only constant in business today is that change is inevitable. This change can manifest in many forms in a business such as

- New system implementations (ERP, CRM, Loyalty, HR, etc.)
- Large system upgrades (ERP, CRM, HR, etc.)
- Mergers & acquisitions
- Divestitures
- Migrating to the cloud
- A new (or expanded) master data management initiative
- A new (or expanded) business intelligence/data warehousing project
- A big data initiative
- A multi-channel program
- System rationalization/retirement
- Responding to new regulations
- New management

The common thread across these disparate initiatives is a strong need for data migration. Typically when a business person thinks about data migration, they think, *This sounds like an IT problem*, and they move on.

**The reality is that without strong business support and a proven approach for the data migration associated with the above mentioned initiatives, you have a higher probability of failure than of success.**

## Did you know?

**More than 80% of data migration projects run over time and/or over budget. Cost overruns average 30%. Time overruns average 41%.**

Bloor Group

**83% of data migration projects either fail or exceed their budgets and schedules.**

Gartner

## Why is there such a high failure rate for projects involving data migration?

The fact that over 80% of projects involving data migration fail is a sobering statistic. What we can be sure about is that none of the organizations that embarked on an initiative involving data migration expected to fail. So the very first question that jumps to mind then is, **WHY SUCH A HIGH FAILURE RATE?**

The unfortunate reality is that the data migration process in support of a larger initiative suffers from a crippling series of misconceptions, challenges, and complexities—all of which contribute to organizations underestimating the effort required to be successful. Typically, the effort associated with successful data migration is underestimated for a variety of reasons:

- **Misconceptions and poor assumptions**
  - Thinking that it's just about moving data from A to B.
  - Assuming that the existing data will fit into the new system.
  - Limited understanding of the true quality (or lack of quality) of the data and the associated effort required to make the data "fit for purpose."
  - Focusing exclusively on the business process that needs to be modeled rather than thinking about the data needed to support it.
  - Assuming that documentation about the current state landscape is complete, accurate, and well-understood.
  - Thinking that data migration is a one-time event done by IT that does not warrant much special attention from business.
- **Risk not shifted early enough**
  - As any good project manager will attest to, the earlier that risk is uncovered in a project, the better. Unfortunately, in a data migration effort, risk usually surfaces very late in the form of load failures in the target system.
  - Often, these load failures are a result of poor data quality and a poor understanding of the nuances contained in the data and how those nuances would impact the load.
- **Lack of formal data governance, methodology, and tools to support the migration**
  - Being able to load the data into the target system is not the measure of success because the data may be valid but not correct.
  - Without a formal approach to data governance, you will have a moving target in terms of trying to figure out how "good" the data should be. This, combined with no clear rules of engagement for the remediation of data quality issues, leads to data that is not "fit for purpose."
  - A traditional waterfall approach to a data migration project does not fit well with the iterative approach required to get the data to the required level of quality.
  - Lack of purpose-built data quality technology to assist in data profiling and cleansing means that organizations don't have a proper understanding of the current state of the data early enough in the project nor the ability to properly clean it.
- **Complex environments**
  - Today's environments are far from simple, and often, the complexity of the environment overshadows the special attention that must be paid to the data to ensure that the target system can deliver on the expected ROI.

# What does it take to be successful when migrating data?



*"The support MRP received from LumenData helped us build a robust repeatable data quality migration process that greatly improved the information used to drive our core business processes."*

*CIO – Market Resource Partners*

A high failure rate does not mean that your data migration efforts are doomed from the start. The following are some clear guidelines for migration success:

1. **Engage the business.** Data migration is not solely IT's problem. Participation from business stakeholders is vital to the success of the migration. Everyone is fully committed in their day jobs, and even though it will be painful, an organization must make these key resources available during the migration process. Failure to make business resources available for current state and to-be discussions will lead to an IT-centric project that involves guesswork and assumptions—most of which will remain hidden until very late in the project when the data fails to load properly or when functional testing fails. Ongoing business involvement will be needed in the form of data remediation for data quality exceptions. How much involvement will depend on the quality of the data and number of exceptions that need to be addressed. Data profiling will provide visibility to the remediation needs.
2. **Create a data migration "factory."** Don't think about the data migration project as a one-off event involving shifting data from A to B. Instead, think of setting up a "factory" process whereby you create an end-to-end, repeatable process flow from source(s) to target with an associated methodology and purpose-built data quality engine at the core.
3. **Profile the data.** The best approach to uncovering risk early and fully understanding your data is through data profiling. The data profiling results can be used to engage stakeholders in aspects such as:
  - a. Empirical discussions about the true state of the data.
  - b. Root cause analysis that can point to business process flaws manifesting in the data.
  - c. Discovering and understanding discrepancies between the documented state vs. the true state of the environment based on the data.
  - d. Engaging in discussions about how "good" the data needs to be, including prioritization of mandatory vs. 'nice to have' requirements.
  - e. Understanding the level of effort required to make the data "fit for purpose."

Without these types of discussions and associated analyses, any estimate about the actual time and effort required to successfully migrate the data is simply guesswork.

4. **Establish data governance.** Data governance is about the people and process aspects of the migration and is as important as any technology being used. Who needs to be involved? How "good" must the data be? What constitutes a duplicate? What should the valid list of values be for particular attributes? What process should be followed to remediate a data quality exception? How should issues be escalated if a decision cannot be made? Who breaks a deadlock? How should you deal with the initial spike of remediation needs vs. steady state? How should the quality of the data be managed in the target system after go-live? These and more are questions that are the responsibility of data governance to manage.
5. **Iterate and accept change.** Each iteration of the flow through the data migration factory reduces risk and improves the quality of the final output through improved visibility into the data and the introduction of additional cleansing rules per iteration. Iterate as many times as possible and then do one more! Realize that change will be a constant bedfellow during a migration, so embrace it.
6. **Separate the data from the application.** Make the distinction that the data and the application are two separate but connected things that have very separate needs. Once you make that leap, then the data migration factory is within your grasp.



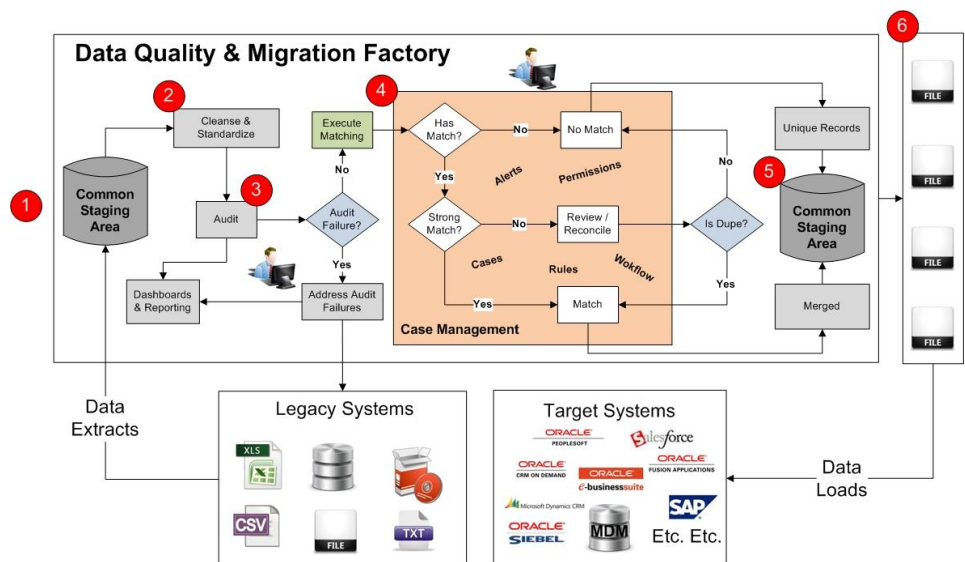
# What does the migration factory look like?



The migration factory approach provides a proven platform and process flow that allows you to separate the data from the application. From here, you can manage the churn in the data and iterate your way to a successful migration. The following capabilities are provided with the data migration factory:

1. Data are extracted from the source system(s) using any data movement technology and loaded into a common staging area. This step forms the entry point into the data migration factory.
2. Cleansing and standardization rules (based on the initial profiling results and user workshops) are executed to address data quality exceptions.
3. The cleansing results are audited to measure the incremental improvement of the quality against specified targets per attribute (or group of attributes). Audit results are available in dashboards and reports. Records that fail the audit are investigated to determine root cause. Additional rules can be introduced in the migration factory to address the exceptions; cleanups can occur in the source system(s) themselves or decisions would need to be made about whether the records should be migrated in the first place if they fail the audit rules.
4. Potential duplicates are identified via a combination of best-in-class exact and fuzzy matching rules. Potential duplicates are routed via a case management process to different users for remediation, or the factory can automatically consolidate those records that don't need manual remediation.
5. Unique records and merged records are outputted to a common staging area where various output files are created based on the specific requirements of the target system(s).
6. Data can then be loaded into the target system(s) using any data movement technology.

The entire process flow is shown below for reference:



The migration factory shown above is based on Oracle's Enterprise Data Quality application, a best-in-class data quality platform. The data migration factory can be run on-premise or in the cloud.

## What are the benefits of adopting this approach for data migrations?

The benefits of adopting a data migration factory approach can be found in various areas:

1. **Risk reduction.** By separating the data from the application, you can guarantee that enough attention is being paid to the often overlooked data architecture and remediation aspects associated of the project.
2. **Improved user adoption.** By allowing the implementation partner for the application to focus more on the user experience vs. having to constantly wrestle with data, a better product will be delivered that more closely aligns with what users actually want. The data feeding the new application will be fit for purpose due to the purpose-built nature of the data migration factory. Given that the #1 reason for CRM failure in the past has been due to poor data, the importance of clean data when it comes to user adoption cannot be overemphasized.
3. **Lower long term costs.** While it may at first seem counter-intuitive to introduce yet another party into your project, recall that more than 80% of projects involving data migration fail or run over budget. Because the problems associated with data migration usually occur very late in a project's lifecycle, organizations tend to throw everything at the project in the hopes of salvaging it. When they look again, the project costs and timeline have spiraled out of control. Adding extra budget up-front to your project to allow you to create the data migration factory will ensure that you don't fall into the trap that more than 80% of organizations fall into.
4. **Increased speed to market.** Organizations don't embark on new projects just for fun. There is a tangible business need behind every new project, and the sooner that project can start delivering benefits, the better. By introducing the data migration concept, you can ensure that your new projects can be delivered on or ahead of schedule; otherwise, you risk becoming part of the 80+% of organizations that exceed their timeline and/or budget when rolling out new projects involving data migration.

## About LumenData

info@lumendata.com  
www.lumendata.com  
+1 (855) MYLUMEN

LumenData is a leading provider of Enterprise Information Management solutions with deep expertise in Master Data Management, Data Strategy, Data Quality, Data Governance, Data Integration, and Big Data.

Through a combination of highly-trained consultants, strong partnerships, relentless focus on quality and executive oversight, LumenData has successfully delivered planning, implementation, integration, maintenance, and training services to over 50 blue chip clients in industries including Financial Services, Higher Education, Life Sciences, Manufacturing, Retail, Telecom, and more.

LumenData's expertise also encompasses the following products and categories:

- MDM – Customer Mastering, Product Mastering (PIM), Hierarchy Management, and Supplier Mastering
- Secure Cloud MDM services include Hybrid MDM and Cloud Integrations
- Customer and Product Data Quality
- Data Strategy and Solution Center services
- Data Services – D&B, Informatica AddressDoctor, Loqate, and Acxiom
- Data Integration using industry-leading middleware systems
- Company-wide data governance, stewardship, and quality metrics

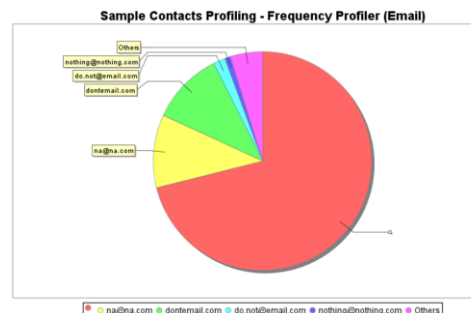
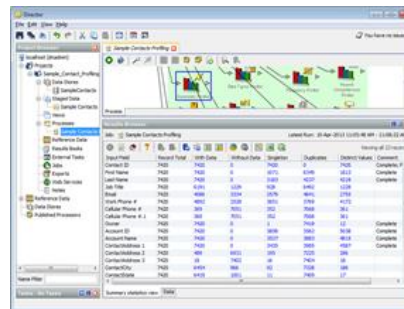
# About Oracle Enterprise Data Quality (OEDQ)

Oracle Enterprise Data Quality (OEDQ) helps organizations achieve maximum value from their business-critical applications by delivering fit-for-purpose data. OEDQ enables individuals and collaborative teams to quickly and easily identify and resolve any problems in underlying data. With OEDQ, customers can identify new opportunities, improve operational efficiency, and more efficiently comply with industry or governmental regulations.

Quick to deploy and easy to use, OEDQ products bring the ability to enhance the quality of data to all stakeholders in any data management initiative. Oracle Enterprise Data Quality provides a business-oriented user interface that offers the following capabilities:

- Profiling, Audit, and Dashboards
- Parsing and Standardization
- Match and Merge
- Case Management
- Address Verification
- Product Data Capabilities

Shown below are some screenshots of the OEDQ user interface:



[www.oracle.com/us/products/middleware/data-integration/enterprise-data-quality/overview/index.html](http://www.oracle.com/us/products/middleware/data-integration/enterprise-data-quality/overview/index.html)