Data Migration Checklist for Legacy System Retirement

A DISCUSSION OF BEST PRACTICES IN LEGACY DATA MIGRATION AND CONVERSION.

CHECKLIST

September, 2020

COPERROR CORE SERVICES

GAINE

THE IMPORTANCE OF LEGACY DATA MIGRATION

As healthcare and life sciences organizations undertake the transformation of their outdated core systems, it is crucial to pay sufficient attention to the migration and conversion of legacy data into the new applications and analytic platforms. The overwhelming majority of projects that involve the migration/conversion of legacy data to a new platform are plagued by costly overruns or even project failures. Research by Gartner confirms the challenges of data migration but the quantum is astonishing.

83%

83% of data migration projects fail or run substantially over budget.

- GARTNER

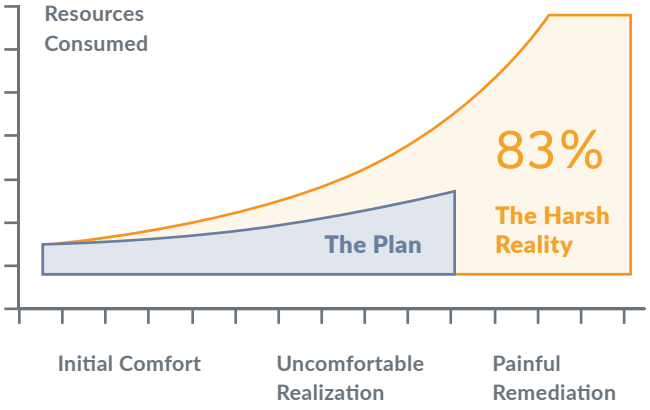
In this paper, we discuss the most important considerations of data migration projects in order to avoid being part of the unfortunate 83% majority.

For more information on critical data migration considerations, visit:
gainesolutions.com/datamigration

FACING REALITY

The harsh reality for the majority of legacy system retirement projects only hits home when the project team realizes that they have underestimated the complexity of converting data and they are faced with an extended period of data remediation.

Commonly, the costs of remediation substantially exceed the original data conversion budget as resources are thrown at the problem in attempts to reach the go-live date through brute force.



What is certain is that none of the 83% believed they would fall victim to the "harsh reality" when they were planning their projects. The data migration effort is underestimated for a variety of reasons:

FACING REALITY

Lack of Data Knowledge

Documentation of legacy systems is typically inaccurate, incomplete or missing entirely. An organization rarely has the people with the time and knowledge of existing data to fill these knowledge gaps.

Data Quality Problems

Even organizations that are aware of the extent of their data quality challenges fail to understand the effort required to fix all that is incomplete, inaccurate or inconsistent in their information.

Delaying Functional Testing

Getting data loaded into a new platform is not a measure of success. The data loaded into new applications may be valid but not correct which only becomes apparent during functional testing. Functional testing should happen early in the plan and continue in parallel with data migration.

Lack of Flexibility & Specification Changes

Gaps in data specifications and inaccurate assumptions of data quality/availability, combine with changes to business requirements to create a "moving target." This in turn creates a huge volume of specification changes. Traditional data migration methods, based on a waterfall project approach, are ill suited to the churn of a data migration effort.

Data Validation & Audit

Validating the business data post-transformation extends beyond merely looking at data in Excel spreadsheets. The data validation process must combine legacy data, the configuration specification and data generated in the new application to provide checksums, record counts and like-for-like comparisons between the old and the new data models. Balancing and validation of financial and statistical data can be a significant undertaking.

THE CHECKLIST

The problems that create project overruns are typically only visible in the later stages of a project when it is too late or very costly to resolve.

Organizations should carefully assess their preparedness prior to starting their data migration in order to avoid the pitfalls. The proper project management attention needs to be given to the data migration/conversion effort for these projects.

Here we discuss eight key areas that should be given close attention before embarking upon a data migration project.

DATA MIGRATION CHECKLIST

1. BUSINESS STAKEHOLDER PARTICIPATION

- [ ] *Identify the business SME's by name and role.*
- [ ] *Agree on business ownership of the data with key business stakeholders.*
- [ ] *Secure dedicated time commitments from SME's throughout the project duration.*
- [ ] *Address and identify the entire data migration / conversion effort as part of the overall project plan.*

PG. 4

2. ESTIMATING EFFORT

- [ ] *Complete the "to-be" model and basic legacy data mapping.*
- [ ] *Profile legacy data against the new model.*
- [ ] *Identify gaps and dependencies in legacy data in collaboration with business SME's.*
- [ ] *Determine what data is to be migrated / converted or archived.*

PG. 4

3. DATA VALIDATION PLAN

- [ ] *Define a process by which data may be validated by SME's.*
- [ ] *Define metrics by which records loaded into the new system may be reconciled to the old system.*

PG. 4

4. MANAGING THE SPECIFICATION

- [ ] *Define a structured process to store and manage changes to the specification.*
- [ ] *Provide a method to report on the impact of changes to the specification.*
- [ ] *Create a link between the specification and the validation reports to test the spec.*

PG. 5

5. REFRESH AND REGRESSION TESTING

- [ ] *Create a process to apply all data manipulations including manual updates / additions.*
- [ ] *Design a set of reconciliation reports for regression testing after a refresh.*
- [ ] *Recognize that conversion/migration testing processes are not linear, but iterative.*

PG. 5

6. GO-LIVE STRATEGY

- [ ] *Define the process to synchronize systems during any post go-live pilot phase.*
- [ ] *Get agreement on what defines the end of the pilot phase.*
- [ ] *Provide for the proper hand-off and education to the support/maintenance team.*

PG. 6

7. ONGOING DATA GOVERNANCE

- [ ] *Include ongoing data governance as part of the core system transformation plan.*
- [ ] *The data governance committee needs to include all business functions as well as IT and have their roles and responsibilities outlined and understood.*

PG. 6

8. PROCESS ABOVE MECHANICS

- [ ] *Ensure that you have a process-based strategy not a tools-based strategy.*
- [ ] *Proper project management must be in place and include data migration / conversion.*

PG. 6

GAINE

1. BUSINESS STAKEHOLDER PARTICIPATION

An organization should be realistic about the time commitment required from business subject matter experts (SME's). The very people that are indispensable in day-to-day business operations are typically the same employees required throughout the system transformation effort for important decisions and key insights.

As painful as it is, an organization must be prepared to make key resources available to the project team throughout the migration effort. Failing to secure the business resources required to define, clarify and validate the conversion will result in the technical team having to guess at business decisions. This guesswork by the project team typically remains hidden until the impact is felt during functional testing in the later stages of the project.

2. ESTIMATING EFFORT

The data transformation effort is driven by complexity more than volume. Basic measures such as the number of data objects or number of records to be converted are not sole indicators of the time and effort required for migration.

Estimates based on hours-per-transformation become meaningless when a single issue in a complex business process may give rise to hundreds of hours of additional effort. Even simple mappings can be complicated by obscure data issues relating to granularity, integrity or consistency of the legacy data.

During the scoping and requirements phase and most definitively before committing to a go-live date, an organization should undertake a data profiling exercise to assess the "as-is" data against the "to-be" model. These insights will help the project team make a more realistic and accurate plan for the conversion effort.

Organizations should also take this opportunity to strategically examine what data should be migrated to the new system. It may be more efficient to archive certain data into a data warehouse rather than migrate this data. Such data would not be relevant for day-to-day processing nor current transactional needs.

3. DATA VALIDATION PLAN

The project team should have a detailed plan for how the new data will be validated against the specification. It is insufficient to say that users will validate data in Excel reports without considering the inventory of reports that will be required.

This inventory of reports and the data required to populate these reports will drive the design of the underlying data store to capture the key metrics and transformation steps throughout the process.

It is not uncommon for a project team involved in a system transformation initiative to produce upwards of 200 validation reports during the project testing phase. Many of these reports rely upon interim counts and temporary calculations. A plan to capture this process information is critical as discovering it during validation is problematic.

4. MANAGING THE SPECIFICATION

Throughout the course of a system transformation project the data conversion specification may undergo hundreds of changes – both large and small. These changes are driven by evolving business requirements, dealing with unknown issues in legacy data, and managing newly identified use-cases that arise during functional testing.

It is not sufficient to simply control updates to the specification documents – although for many this is a challenge in itself. You should be able to determine the impact of each change to the specification so that the number of transformations or data objects affected by each change is understood. This helps plan how to deal with the change and how to manage the impact upon the project plan.

Another consideration is that within a "to-be" data specification, a single data element may be included within multiple areas, and therefore a change in one area may cause unforeseen problems in another. It is extremely difficult, if not impossible, to reliably manage these interdependencies using a manual process and spreadsheets.

A final thought on managing the specification is that each requirement expressed within the specification should be linked to one or more validation reports. When the specification changes, so should the associated validation process.

The specification should be tightly managed, preferably within a database where the integrity can be enforced and changes can be more easily analyzed. Remember, the specification is not just documentation artifact; it is the central point of control throughout the project.

5. REFRESH AND REGRESSION TESTING

At some point in the migration project, the source data will need to be refreshed. Typically data is refreshed prior to each testing cycle and certainly prior to go-live. All phases of migration testing feeds back to previous phases, which can trigger changes throughout. This requires testing be done using an iterative process until each part of the migration/conversion has successfully passed the testing stage.

Two challenges present themselves during a refresh. The first is to ensure that all of the transformations and data manipulations applied to-date are re-applied in the correct sequence. The second is to validate that the refresh has not introduced (or re-introduced) any data errors.

The ability to reload the data and to test each requirement in the specification against the results is considered a test harness and is a critical element of the migration approach.

6. GO-LIVE STRATEGY

Whether as part of a phased implementation or a "big bang" go-live strategy, it is important to plan for business operations during and immediately following the period of go-live.

There is inevitably a period of time between the final refresh of legacy data and completing the load and validation of the new platform. In a large project, this period may stretch into several weeks. During this time either the legacy systems are "frozen" and no new updates are allowed, or a process to capture these legacy updates and apply them to the new application platform is required.

Project managers should involve the support team in the design and development of the new system and not just turn over the data migration/conversion to the team after the pilot implementation occurs. The support or maintenance group, generally overloaded with regulatory system changes and system enhancements, needs to be properly educated and involved in the data migration/conversion.

Post go-live, attention quickly shifts to the auditors who are required to ensure that the data loaded into the new systems can be reconciled with the retired legacy systems. Auditors use a combination of high-level aggregations and specific detailed tests to satisfy their analysis. The migration team should not overlook the need to collect detailed history and lineage of all the new data to support this audit process. If this detailed audit trail is not readily available, then the audit process will become an unwelcome distraction to the business as it adjusts to the use of the new application.

7. ONGOING DATA GOVERNANCE

The transformation of legacy systems presents an ideal opportunity to introduce formal data governance processes to an organization. The data migration project will require the collaboration of IT and business stakeholders to define data rules and policies. These same rules and policies may be used as the initial basis for data governance within the organization.

No application system is infallible; at the root cause of bad data are poor process controls and a lack of standards and training for system users. These same underlying factors will continue to introduce errors in the new system and dilute the investment in data quality throughout the migration.

8. PROCESS ABOVE MECHANICS

Be wary of an over-emphasis of the ability to manipulate data and connect systems rather than the ability to manage the migration process.

Technical staff may present a compelling plan for the transformation of legacy data but the mechanics of the transformation are just a small part of the overall recipe for success.

Concentrate the project planning on the first seven points in this section before you become concerned with the tool that handle database connections and transformations.

COPEROR CORE MDM PLATFORM OVERVIEW

It is no coincidence that the Coperor data migration platform developed by Gaine Healthcare helps tick all the boxes discussed in the previous section.

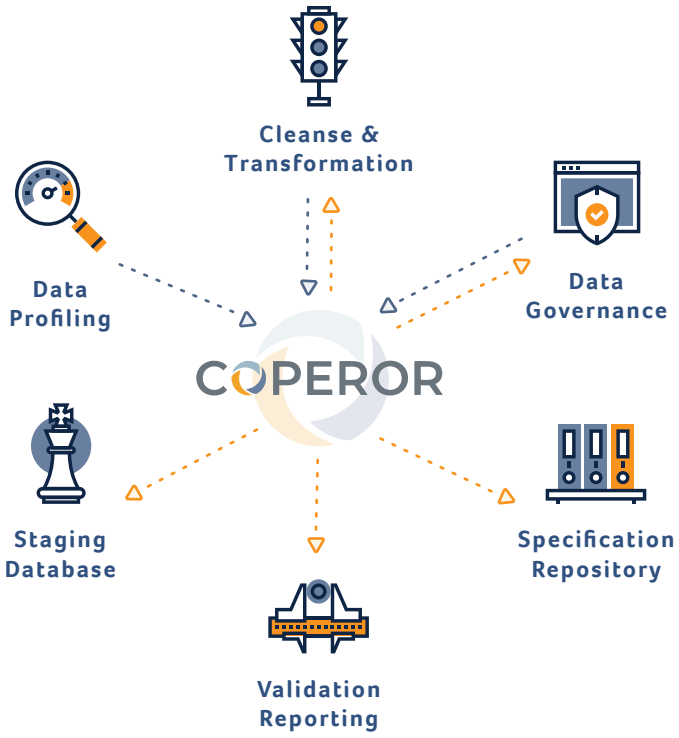
Coperor is a purpose-built platform to address the challenges of legacy data migration. Since 2007 the Coperor platform has been used to retire more than 2,000 legacy applications in support of core system transformation initiatives and continues to help companies manage risk and reduce cost within data migration projects.

Coperor combines multiple components in an integrated environment to support a best-practice methodology for data migration.

Process control is at the heart of the Coperor solution. With Coperor all aspects of the migration can be coordinated and integrated, and throughout the migration no change to either the specification or business data goes unmanaged. This level of control allows data to be refreshed in an automated manner, and it provides a central point of control for the project team.

Coperor provides not only cleansing, transformation and matching capabilities but can also integrate with third-party tools or custom programs if these are available. The Coperor platform provides a data profiling service to assess the legacy data against the new data model as well as a specification process which affords control of the specification, impact analysis and a link to validation reports.

Data governance workflows are provided to enable ongoing management of data quality. Coperor data governance templates include workflows for



managing changes to key attributes on the master record, handling conflicts between matching records, resolving suspect duplicates between systems, reconciling inconsistent relationships between records and many others.

Coperor should not be compared to ETL tools or data cleansing tools. Coperor does not replace the transformation or connectivity provided by an ETL tool, but it does enable an ETL tool to be used in a controlled and effective manner. The Coperor platform provides an integrated environment ideally suited to the churn of legacy data migration without the need to design and build this capability.

Furthermore, it eliminates the need for the user to build out all of these processes thereby saving significant time and money.

SUMMARY

Legacy data migration is more difficult than it initially appears. Getting it wrong is expensive and painful.

Don't let the mechanics of how legacy data will be transformed overshadow the need to manage the process.

Give careful and objective scrutiny to the checklist in section two of this document, and if you have any doubt about your approach be prepared to pause. It is far cheaper and easier to fill a gap up-front than to dig your way out of trouble later in the project.

Data migration is a specialist discipline and requires the coordination of a lot of moving parts. It's not enough to say "we have smart people and good tools." You need experienced people and integrated tools to be successful.

The Coperor platform is much more than just an accelerator to a legacy migration project. While Coperor does relieve the project team from the overhead of designing and building a migration capability, it more importantly encourages and enforces a best-practices approach.

The project team will find that each component is seamlessly integrated with the next, and very soon everything from upfront data profiling to data validation during testing and ongoing data governance work together to provide visibility and control to all stakeholders.

ABOUT GAINE

Gaine helps leading healthcare organizations, life sciences, and biotech companies leverage information as they strive for competitive advantage and operational efficiency. Gaine Healthcare's Coperor™ platform arms its clients with Information Agility™ which maximizes flexibility, minimizes cost and time-to-value, and eliminates the risks associated with implementing new systems, migrating existing systems, and integrating data from internal and external sources.

3889 Long St., Ste 201
San Luis Obispo, CA 93401

(888) 403-1301
sales@gainehealthcare.com

gaine.health