Experian™
Data Quality

Data Migration Project Checklist

Experian Pandora

Created by Dylan Jones
Editor Data Migration Pro.com

Contents

1. Introduction

Data Migration projects are notorious for high failure rates but what is the principle cause of failure? The root of the problem lies with a lack of awareness. Data Migration sponsors and project leaders simply underestimate the many activities that are required throughout the data migration lifecycle.

Project leaders often ignore key requirements such as the need for adopting a rigorous methodology, the right technology and robust data quality principles. Omissions like these cost organisations heavily in the form of unplanned project delays and costs, not to mention the frustration from business leaders as the planned go-live date for the target system slips by months, often years.

This document will help data migration project leaders by arming them with an extensive checklist that can be used to validate their own approach. Think of it as a safety net to ensure that your data migration strategy has the necessary components for success.

I hope you find this checklist a useful tool for improving the success rate of your future data migrations.

2. About Experian Pandora Software

A radically new approach to data migration technology
Experian Pandora provides next generation technology for modern data demands. Focused on the areas of Data Quality Management, Data Migration and Data Governance, our unique solution - Experian Pandora is transforming the way data migration projects are delivered.

Traditional data migration technologies require resources and investment that you simply won't need with Experian Pandora. With Experian Pandora, the need for separate ETL, Data Profiling, Data Quality, Testing and Stewardship software is completely eliminated. What's more, Pandora runs on commodity hardware with users becoming proficient in days.

Want to experience Experian Pandora for yourself?
Download the Experian Pandora free data profiler and start managing your data quality today. When you're ready, feel free to request a demo of our full Experian Pandora Data Migration solution. Alternatively, read a detailed report from Bloor research on why Experian Pandora represents a radically new approach to tackling data migration projects.

- Download Free Pandora Profiler
- Request a Demo of Pandora Data Migration Edition

To find out more about Experian Pandora, please visit www.edq.com/uk

"Using Pandora, an associate analyst completed a data migration impact assessment of an entire government system in less than 3 days. Conventional approaches would have taken weeks." **John Morris, author of Practical Data Migration**

3. Phase 1: Pre-Migration Planning

Have you assessed the viability of your migration with a pre-migration impact assessment?

Most data migration projects go barrelling headlong into the main project without considering whether the migration is viable, how long it will take, what technology it will require and what major obstacles lie ahead. It is therefore advisable to perform a pre-migration impact assessment to verify the cost and likely outcome of the migration. The later you plan on doing this the greater the risk. Adjust your strategy accordingly.

Have you based project estimates on guesswork or a more accurate assessment?

Don't worry, you're not alone, most projects are based on previous project estimates at best or optimistic guesswork at worst.

Once again, your pre-migration impact assessment should provide far more accurate analysis of cost and resource requirements so if you have tight deadlines, a complex migration and limited resources make sure you perform a migration impact assessment asap.

Have you made the business and IT communities aware of the need for their involvement?

It makes perfect sense to inform the relevant data stakeholders and technical teams of their forthcoming commitments before the migration kicks off.

It can be very difficult to drag a subject matter expert out of their day job for a 2-3 hours analysis session once a week if their seniors are not fully aware of what lies ahead.

In addition, by identifying what resources are required in advance you will eliminate the risk of having gaps in your legacy or target skillset.

There are numerous aspects of a data migration that require business sign-off and commitment. Get in front of sponsors and stakeholders well in advance and ensure they understand AND agree to what their involvement will be.

Have you formally agreed the security restrictions for your project?

I have painful memories of one migration where we thought everything was in place so kicked off the project only to be promptly shut down on the very first day.

We had assumed that the security measures agreed with the client project manager were sufficient but we did not account for the corporate security team getting demanding a far more stringent set of controls. This resulted in an unplanned 8 weeks project delay.

Don't make the same mistake.

Obtain a formal agreement from corporate security in advance and follow security policies to the letter. Simply putting your head in the sand and hoping you won't get caught is highly unprofessional and clearly a risk-fuelled strategy.

Have you identified your key project resources and when they are required?

Don't start your project hoping that an agency or systems integrator will magically provision those missing resources you need.

I met a company several months ago who decided they did not require a lead data migration analyst because the 'project plan was so well defined'. Suffice to say they're now heading for major delays as the project spins out of control through lack of specialist leadership.

Make sure you understand precisely what roles are required on a data migration and ensure you have a plan for bringing those roles into the project at the right time.

For example, there is a tendency to launch a project with a full contingent of developers armed with tools at the ready and raring to go. This is both costly and unnecessary. A small bunch of data migration, data quality and business analysts can perform the bulk of the migration discovery, quality analysis and mapping well before the developers get involved, often creating a far more successful (and much cheaper) overall migration.

The lesson is to understand the key migration activities and dependencies then plan to have the right resources available when required.

Have you determined the optimal project delivery structure?

Data migrations do not always suit a waterfall approach yet the vast majority of data migration plans I have witnessed nearly always resemble a classic waterfall design of Analyse, Design, Build, Test and Launch, all executed in a linear fashion.

Agile, iterative project planning, with tangible delivery drops are often far more effective in creating value for the business so ensure that your overall plan is flexible enough to cope with the likely change events that will occur.

Have you created a structured task workflow so each member will understand what tasks are expected and in which sequence?

Most data migration project plans consist of some vague drop dates or timelines that indicate when the business or technical teams require a specific release or activity to be completed. What this will not show you is the precise workflow that will get you to those points. This needs to be ideally defined before project inception so that there is no confusion as you move into the initiation phase. Carrying this out before project inception will also help you identify gaps in your resourcing model where the necessary skills or budgets are lacking.

Have you created the appropriate training documentation and designed a training plan?

Data migration projects typically require a lot of additional tools and project support platforms to function smoothly so you must ensure that all your training materials and education tools are tested and in place prior to project inception.

Ideally you would want all the resources to be fully trained in advance of the project but if this isn't possible at least ensure that training and education is factored into the plan.

Do you have a configuration management policy and the appropriate software in place?

Data migration projects create a lot of resource materials. Profiling results, data quality issues, mapping specifications, interface specifications - the list is endless. The sheer volume of materials required on a project can soon overwhelm team members and finding the latest, most accurate versions can be extremely difficult without a policy in place.

Ensure that you have a well-defined configuration management approach in place before project inception. You don't want to be stumbling through project initiation trying to test out different approaches.

Have you planned for a secure, collaborative working environment to be in place?

If your project is likely to involve 3rd parties and cross-organisational support then it pays to use a dedicated product for managing all the communications, materials, planning and coordination on the project. Having a unified project 'hub'

will also make your project run smoother if this is configured and ready prior to project initiation.

Have you created an agreed set of data migration policy documents?

How will project staff be expected to handle data securely? Who will be responsible for signing off data quality rules? What escalation procedures will be in place?

There are a multitude of different policies required for a typical migration to run smoothly. It pays to agree these in advance of the migration so that the project initiation phase runs effortlessly.

4. Phase 2: Project Initiation

Have you created a stakeholder communication plan and stakeholder register?

During this phase you need to formalise how each stakeholder will be informed. We may well have created an overall policy beforehand but now we need to instantiate it with each individual stakeholder.

Don't create an anxiety gap in your project where your stakeholders witness no discernible progress for months at a time.

Determine what level of reporting you will deliver for each type of stakeholder and get agreement with them on the format and frequency. Dropping them an email six months into the project explaining why you're headed for a 8 week delay will not win you any sympathy from poorly engaged business leaders.

To communicate with stakeholders obviously assumes you know who they are and how to contact them so record all the different types of stakeholder and their preferred communication pathway.

Have you tweaked and published your project policies?

Now is the time to get your policies completed and circulated across the team and new recruits. Any policies that define how the business will be involved during the project also need to be circulated and signed off.

Don't assume that everyone knows what is expected of them so get people used to learning about and signing off project policies early in the lifecycle.

Have you created a high-level, first-cut project plan?

If you have followed best practice and implemented a pre-migration impact assessment you should have a reasonable level of detail for your project plan. If not, simply complete with as much detail as possible with an agreed caveat that the data discovery exercises will drive the strategy of the project.

Tip: It is still beneficial to carry out a migration impact assessment during the

initiation phase irrespective of the analysis activities that take place in the next phase.

You cannot create accurate timelines for your project plan until you have analysed the data.

For example, simply creating an arbitrary 8-week window for data cleansing activities on your project plan is meaningless if the data is found to be truly abysmal.

It is also vital that you understand the dependencies in a data migration project in some considerable depth. You can't build the mappings in a data migration tool until you have discovered the relationships and you can't do this until the analysis and discovery phase has completed.

Don't just rely on a carbon copy of a previous data migration project plan or some template downloaded from the web. Your plan will be dictated by the conditions found on the ground and the wider programme commitments that your particular project dictates.

Have you created your standard project document templates?

During this phase you must create project documentation such as risk register, issue register, acceptance criteria, project controls, job descriptions, project progress report, change management report, RACI etc.

They do not need to be complete but they do need to be formalised with a process that everyone is aware of.

Have you defined and formalised your 3rd Party supplier agreements and requirements?

Project initiation is a great starting point to determine what additional expertise is required but ideally this should have been done prior to the project commencing.

Don't leave assumptions when engaging with external suppliers, there should be clear instructions on what exactly needs to be delivered. One particular grey area that many project leaders ignore is the agreement surrounding data quality. Who will be responsible for correcting data defects on your project? Make sure this is clearly defined in all supplier agreements.

Have you scheduled your next phase tasks adequately?

At this phase you should be meticulously planning your next phase activities to ensure that the business and IT communities are aware of the workshops they will be involved in and the deliverables they will be expected to create.

Have you resolved any security issues and gained approved access to the legacy datasets?

Don't assume that because your project has been signed off you will automatically be granted access to the data. Get approvals from security representatives (before this phase if possible) and consult with IT on how you will be able to analyse the legacy and source systems without impacting the business.

Full extracts of data on a secure, independent analysis platform is the best option but you may have to compromise.

It is advisable to create a security policy for the project so that everyone is aware of their responsibilities and the professional approach you will be taking on the project. Designate someone as the 'security officer' on the project so that there is a single point of contact regarding security queries.

Have you defined the hardware and software requirements for the later phases?

What machines will the team run on? What software will they need? What licenses will you require at each phase?

Sounds obvious but not for one recent project manager who completely forgot to put the order in and had to watch 7 members of his team sitting idly by as the purchase order crawled through procurement!

Don't make the same mistake. Look at each phase of the project and determine what software and hardware will be required.

For example, to complete your project successfully you will require combinations of the following:

- Data discovery and profiling software
- Data quality management, cleansing and monitoring software
- ETL/ data movement software
- Data stewardship and testing software

You will also need to determine what operating systems, hardware and licensing is required to build your analysis, test, QA and production servers. It can often take weeks to procure this kind of equipment so you ideally need to have done this even before project initiation.

5. Phase 3: Landscape Analysis

Have you created a detailed data dictionary?

A data dictionary can mean many things to many people but it is advisable to create a simple catalogue of all the information you have retrieved on the data under assessment. Make this tool easy to search, accessible but with role-based security in place where required. A project wiki is a useful tool in this respect.

Have you created a high-level source to target mapping specification?

At this stage you will not have a complete source-to-target specification but you should have identified the high-level objects and relationships that will be linked during the migration. These will be further analysed in the later design phase.

Has the risk management process been shared with the team and are they continuously updating the risk register as issues are discovered?

There will be many issues discovered during this phase so make it easy for risks to be recorded. For example, create a simple online form where anyone can add risks during their analysis, you can also filter them out later but for now we need to gather as many as possible and see where any major issues are coming from.

You will need to create a separate data quality risk management process to cope with the large volume of issues discovered. Manage this proactively in weekly workshops with a cross-section of business, IT and project staff to ensure rapid resolution of the problems discovered.

Have you created a data quality management process and impact report?

If you've been following our advice on Data Migration Pro.com and Data Quality Pro.com you will know that without a robust data quality rules management process your project will almost certainly fail or experience delays.

Understand the concept of data quality rules discovery, management and resolution so you deliver a migration that is fit for purpose.

The data quality process is not a one-stop effort, it will continue throughout the project but at this phase one of our main goals is discovering the impact of the data so decisions can be made that could affect project timescales, deliverables, budget and resourcing.

Carry out your data quality management process through the use of high performance data discovery, profiling and data quality management software to find, measure, improve and monitor the many thousands of data quality rules you need to govern on the project.

Have you created and shared a first-cut system retirement strategy?

Now is the time to begin warming up the business to the fact that their current systems will be decommissioned post-migration. Ensure that stakeholders are briefed on the aims of the project and start the process of discovering what is required to terminate the legacy systems. Better to address this now than to leave it until later in the project when politics may prevent progress.

Have you created conceptual/logical/physical and common models?

These models are incredibly important for communicating and defining the structure of the legacy and target environments plus the common model that will bridge them together.

The reason we need so many modelling layers is to understand all aspects of the migration from the deeply technical through to how the business community run operations today and how they wish to run operations in the future.

We will be discussing the project with various business and IT groups so the different models help us to convey meaning for the appropriate community.

Creating conceptual and logical models also help us to identify gaps in thinking or design between the source and target environments far earlier in the project so we can make corrections to the solution design.

Have you refined your project estimates?

Most projects start with some vague notion of how long each phase will take. Use your landscape analysis phase to determine the likely timescales based on data quality, complexity, resources available, technology constraints and a host of other factors that will help you determine how to estimate the project timelines.

6. Phase 4: Solution Design

Have you created a detailed mapping design specification?

By the end of this phase you should have a thorough specification of how the source and target objects will be mapped, down to attribute level. This needs to be at a sufficient level to be passed to a developer for implementation in a data migration tool.

Note that we do not progress immediately into build following landscape analysis. It is far more cost-effective to map out the migration using specifications as opposed to coding which can prove expensive and more complex to re-design if issues are discovered.

Have you created an interface design specification?

At the end of this stage you should have a firm design for any interface designs that are required to extract the data from your legacy systems or to load the data into the target systems.

For example, some migrations require change data capture functionality so this needs to be designed and prototyped during this phase.

Have you created a data quality management specification?

This will define how you plan to manage the various data quality issues discovered during the landscape analysis phase. These may fall into certain categories such as:

- Ignore
- Cleanse in source
- Cleanse in staging process
- Cleanse in-flight using coding logic
- Cleanse on target

Have you defined your production hardware requirements?

At this stage you should have a much firmer idea of what technology will be required in the production environment.

The volumetric and interface throughput performance should also be known as you need to specify the appropriate equipment, RAID configurations, operating systems etc.

Have you agreed the service level agreements for the migration?

At this phase it is advisable to agree with the business sponsors what your migration will deliver, by when and to what quality.

Quality, cost and time are variables that need to be agreed upon prior to the build phase so ensure that your sponsors are aware of the design limitations of the migration and exactly what that will mean to the business services they plan to launch on the target platform.

7. Phase 5: Build & Test

Has your build team documented the migration logic?

The team managing the migration execution may not be the team responsible for coding the migration logic. It is therefore essential that the transformations and rules that were used to map the legacy and target environments are accurately published. This will allow the execution team to analyse the root-cause of any subsequent issues discovered.

Have you tested the migration with a mirror of the live environment?

It is advisable to test the migration with data from the production environment, not a smaller sample set. By limiting your test data sample you will almost certainly run into conditions within the live data that cause a defect in your migration at runtime.

Have you developed a reconciliation testing strategy?

Many projects base the success of migration on how many "fall-outs" they witness during the process. This is typically where an item of data cannot be migrated due to some constraint or rule violation in the target or transformation data stores. They then go on to resolve these fall-outs and when no more loading issues are found carry out some basic volumetric testing.

"We had 10,000 customers in our legacy system and we now have 10,000 customers in our target, job done".

Case Study: I recently took a call from a community member based in Oman. Their hospital had subcontracted a data migration to a company who had since completed the project. Several months after the migration project they discovered that many thousands of patients now had incomplete records, missing attributes and generally sub-standard data quality.

Never underestimate the importance of reconciliation testing.

Reconciliation testing is vital to ensure that all of your legacy data has migrated successfully. Create a robust testing process that can guarantee the quality of the data migrated compared to the original legacy data.

Have you defined your reporting strategy and associated technology?

Following on from the previous point, you need to create a robust reporting strategy so that the various roles involved in the project execution can see progress in a format that suits them.

For example, a data migration manager may wish to see daily statistics whereas a data migration operator will prefer to see runtime statistics. Business sponsors may wish to see weekly performance but business users may wish to progress by business function or even data subject areas.

If you have created service level agreements these need to be incorporated into the reporting strategy so that you can track and verify progress against each SLA.

Have you defined an ongoing data quality management solution?

Data quality is continuous and it should certainly not cease when the migration has been delivered. The new users of the system may well introduce errors through inexperience so plan for this by building an ongoing data quality management environment for the target platform.

Ideally, you should use the same data quality technology deployed on the data migration as it will be fully populated with all the data quality rules required for ongoing management and maintenance.

Have you created a data migration fall-back policy?

What if the migration fails? How will you rollback? What needs to be done to facilitate this?

The best advice is to hope for the best but plan for the worst. Fall-back planning can often be complex and require cross-organisation support so plan well in advance of the go-live date.

Have you confirmed and planned your legacy decommission strategy?

By now you should have a clear approach, with full agreement, of how you will decommission the legacy environment following the migration execution. You need to be communicating with all stakeholders exactly what will happen to their systems after the migration completes.

Have you completed all relevant execution training?

The team running the execution phase may differ to those on the build phase. The migration execution can be complex so ensure that the relevant training materials are planned for and delivered by the end of this phase.

Have you obtained sign-off for anticipated data quality levels in the target?

It is rare that all historical data defects can be resolved prior to migration but at this stage you should certainly know what they are and what impact they will cause.

The data is not the responsibility of the migration load team; it belongs to the business so ensure they sign off any anticipated issues so that they are fully aware of the limitations the data presents.

Have you defined the data migration execution strategy?

Some migration executions can be over in hours, others can take weeks or even months. You will need to create a detailed plan of how the migration execution will take place. This includes sections such as what data will be moved, who will sign-off each phase, what tests will be carried out, what data quality levels are anticipated, when will the business be able to use the data and what transition measures need to be taken.

This can become quite a considerable activity so as ever, plan well in advance.

Have you created a gap-analysis process for measuring actual vs current progress?

This is particularly appropriate on larger scale migrations.

If you have indicated to the business that you will be executing the migration over an 8 week period and that specific deliverables will be created you can then map that out in a chart with time points and anticipated volumes.

As your migration executes you can then chart actual vs estimated so you can identify any gaps or expected delays.

8. Phase 6: Execute & Validate

Have you kept an accurate log of SLA progress?

You will need to demonstrate to the business sponsors and independent auditors that your migration has been compliant. How you will do this varies but if you have agreed SLA's in advance these need to be reported against.

Have you independently validated the migration?

You cannot rely on your migration architecture to validate the migration. An independent process is advisable to ensure that the migration process has delivered the data to a sufficient quality level to support the target services.

9. Phase 7: Decommission & Monitor

Have you completed your system retirement validation?

There will typically be a number of pre-conditions that need to be met before a system can be terminated. Ensure these are fully documented and agreed (this should have been done in earlier phases) so you can begin validating that the migration has met these conditions.

Have you handed over ownership of the data quality management environment?

Close down your project by passing over the process and technology adopted to measure data quality during the project.

Dylan Jones
Editor, Data Quality Pro

Dylan Jones is the Editor and Founder of Data Quality Pro and Data Migration Pro. He has over 20 years' experience of delivering complex data quality and data-driven initiatives. Dylan is a keynote speaker, author and regular publisher of expert insights on data quality related topics.

About Experian Data Quality

Experian Data Quality has built up exceptional market coverage assisting customers with their unique data quality challenges. We provide a comprehensive toolkit for data quality projects combining our market leading software with a vast scope of reference data assets and services. Our mission is to put our customers in a position to make the right decisions from accurate and reliable data. The size and scope of data management projects varies considerably but the common factor in all ventures is unlocking operational efficiency and improving customer engagement. We see the potential of data. Whether it's in enabling ambulances to be sent to the exact location of an emergency or attributing charitable donations to the people who need it the most - data accuracy makes all the difference to service provision.

Experian Australia Pty Ltd
Level 6, 549 St Kilda Road
Melbourne, VIC 3004, Australia

T (61) 3 8699 0100 | F (61) 3 9923 6280
E info@au.experian.com | W edq.com/au

Experian™
Data Quality

Intelligent interactions.
Every time.