

Legacy Data Migration Checklist: Extract Data from Outdated Systems

This checklist is designed for:

Project managers, data engineers, architects and technology leaders responsible for planning and delivering legacy data migrations.

Legacy data migrations are complex by nature. They're often built on old technology and poorly documented, leading to decades of accumulated data issues. Compounding these challenges is the gradual loss of platform expertise, leaving migration teams to reconstruct critical data under time pressure.

This legacy data migration checklist focuses specifically on the challenges of extracting data from ageing systems such as mainframes, AS/400 environments, custom-built applications and discontinued software platforms. It addresses the core decisions every legacy data migration must confront: what data should be migrated, what should be archived, and where a hybrid approach is more appropriate.

This guide also reflects Australian data retention, compliance and decommissioning requirements, so legacy migrations stand up to regulatory scrutiny.

What Makes Legacy Data Migration Different

Legacy data migration introduces challenges that do not typically exist in modern environments. Older systems often come with limited documentation, meaning teams must piece together how data is structured and how systems interact just to understand what they're moving. Outdated technology can add further complications, with ageing platforms and limited integration options restricting how you can approach the migration.

Data quality is another major factor. Over time, duplicate records, inconsistent formats and outdated information naturally build up. On top of that, connections between systems are often poorly understood and only discovered during the migration itself, increasing project delays and complexity.

Critically, not all legacy data should be migrated. Decisions around decommissioning and data archival will have a profound impact on the migration's cost and complexity.

These factors demand specialised migration approaches, which this checklist is designed to address directly.

Your Complete Legacy Data Migration Checklist

This checklist is designed to help teams extract value from legacy systems without carrying unnecessary risk into modern platforms.

Discovery and Reverse Engineering

When documentation is missing or outdated, reverse engineering becomes essential. These discovery tasks help establish a clear understanding of the data landscape before migration begins.

Identify all data sources and databases: Include shadow systems, departmental tools and forgotten environments.

Map data structures and schemas: Use inspection and reverse engineering tools to reconstruct schemas and relationships.

Capture institutional knowledge: Interview long-term employees to document undocumented business context and data usage.

Identify data ownership: Map data owners by domain, even where ownership is informal or unclear.

Document dependencies and integrations: Identify hidden integrations, data flows and upstream or downstream dependencies.

Catalogue custom logic and processes: Record custom code, stored procedures, triggers and batch processes.

Assess actual data volumes: Measure data volumes by source, including historical accumulation and growth patterns.

Identify obsolete or unused data: Flag data that no longer supports active business processes.

Document access patterns: Capture who accesses the data, from which systems, and how frequently.

Map compliance and retention requirements: Document obligations under the Australian Privacy Principles and relevant industry regulations.

Data Quality Assessment

Legacy systems commonly suffer from data quality issues caused by years of inconsistent data entry and limited validation. A comprehensive assessment is critical before any migration effort.

Profile data quality across sources: Use automated profiling tools to identify systemic quality issues.

Identify duplicate records: Detect duplicates within and across systems.

Assess data completeness: Identify missing or partially populated required fields.

Validate data formats: Check formats for dates, phone numbers, addresses and other structured fields.

Check referential integrity: Identify broken relationships and orphaned records.

Identify anomalies and outliers: Flag impossible values, outliers and inconsistencies.

Assess data accuracy: Validate samples against trusted reference sources where available.

Document transformation and cleansing rules: Define required data standardisation and remediation steps.

Estimate your remediation effort: Make a realistic, evidence-based determination of the size of your remediation.

Prioritise remediation by business impact: Focus effort where poor data quality creates the greatest risk.

Australian considerations:

Validate Australian address standards: Confirm correct handling of postcodes and state or territory abbreviations.

Confirm privacy compliance: Ensure handling of personal information aligns with the Australian Privacy Principles.

Validate retention requirements: Confirm data retention periods meet regulatory and industry obligations.

Migration vs Archival

Deciding whether to migrate or archive legacy data is a critical control point in any legacy data migration program. Not all legacy data should be migrated to modern platforms. Given the high costs, complexities and risk of migration, it pays to take a strategic, considered approach. This is how you decide:

Define data retention obligations: Confirm requirements under the Privacy Act 1988 and relevant industry regulations.

Identify active versus dormant data: Classify data based on usage patterns and operational dependency.

Compare migration and archival costs: Evaluate migration effort and platform costs against archival alternatives.

Assess business value of historical data: Determine whether historical data delivers ongoing operational or analytical value.

Design read-only archival solutions: Implement searchable, read-only archives for data that is not migrated.

Plan for legal hold and eDiscovery: Ensure archived data remains accessible for legal, audit and regulatory requests.

Migrate or Archive?

Migrate when data is:

- Actively used
- Required for core business processes
- Legally required to remain online
- Tightly integrated with new systems.

Archive when data is:

- Historical
- Infrequently accessed
- Retained primarily for compliance,
- Costly to migrate relative to its ongoing value.

Hybrid approach: Use when only a subset of data is required for day-to-day operations, while historical records must be retained for compliance, audit or reference purposes without ongoing modification.

Australian considerations:

Validate privacy retention requirements: Ensure alignment with Privacy Act retention and destruction obligations.

Confirm APRA requirements: Address APRA record-keeping and accessibility expectations for regulated entities.

Assess data sovereignty obligations: Ensure regulated data meets Australian data residency and sovereignty requirements.

Data Cleansing and Preparation

Once you identify data quality issues, structured cleansing prepares legacy data for reliable migration into modern systems.

Define data cleansing and transformation rules:

Agree clear rules for how data will be corrected, standardised and transformed.

Deduplicate records using business-approved criteria:

Remove duplicates based on rules validated by business and data owners.

Standardise data formats: Normalise formats for dates, names, addresses and identifiers to ensure consistency.

Populate missing required fields: Apply agreed rules or reference data to complete mandatory fields where possible.

Correct referential integrity issues: Resolve broken relationships and orphaned records before migration.

Validate cleansed data against quality benchmarks:

Confirm accuracy, completeness and consistency meet defined thresholds.

Document all transformations: Maintain traceability for audit, compliance and post-migration review.

Obtain business approval before migration execution:

Secure formal sign-off that cleansed data is fit for migration.

Migration Execution

Executing a legacy data migration requires additional validation due to unreliable or inconsistent source data.

Extract data using appropriate legacy tooling:

Where available, use tools suited to mainframes, AS/400 environments or bespoke legacy platforms.

Apply approved transformation and cleansing rules:

Enforce agreed rules consistently during data extraction and load.

Load data in the correct sequence:

Preserve relationships and dependencies by loading parent and reference data first.

Perform comprehensive data reconciliation:

Validate record counts, key totals and other reconciliation checks between source and target.

Validate business rules in the target system:

Confirm migrated data behaves correctly within downstream applications and processes.

Conduct end-user testing of migrated data:

Involve business users to validate data accuracy and usability.

Document issues and resolutions:

Record defects, remediation actions and known limitations for audit and handover purposes.

Plan for iterative clean-up post go-live:

Allow for controlled remediation cycles after cutover to address residual data quality issues.

Common Legacy Data Migration Challenges and Solutions

Legacy migrations face predictable challenges. Addressing them early helps prevent delays, disruptions and unnecessary effort.

Lack of documentation

Problem: Unknown schemas, missing business rules and undocumented data relationships.

Mitigation:

- Use reverse engineering tools.
- Conduct interviews with long-term staff.
- Engage specialists experienced in legacy environments.

Poor data quality

Problem: Decades of duplicate records, inconsistent formats and missing or invalid values.

Mitigation:

- Apply automated data profiling.
- Structured data cleansing.
- Business-approved quality standards – prioritising critical datasets first.

Obsolete technology

Problem: Unsupported platforms, proprietary formats and limited connectivity options.

Mitigation:

- Use specialist extraction tools, middleware and phased migration approaches supported by deep legacy platform expertise.

Related Data Migration Resources

- ✓ **Data Migration Framework:**
Strategic guidance for complex migration programs
- ✓ **Data Migration Checklist Library:**
Coverage of cloud, database and ERP migrations
- ✓ **Database Migration Checklist:**
For legacy systems involving database modernisation

Modernising legacy systems?

Interactive's data migration services support Australian enterprises with decades of legacy system experience and proven methodologies – backed by prompt, Australian-based support.

[CONTACT US](#)