

# CASE STUDY

## ETL Success Begins With Clean Data

**INDUSTRY**  
Energy

**COUNTRY**  
USA



Organizations periodically encounter the need to replace obsolete enterprise systems as they reach end of life, face the loss of vendor support, or fail to keep up with evolving business requirements. One important step in the replacement process is the mapping and migration of the obsolete system's data to the new system. Ideally, this migration process should be a straightforward extract-transform-load (ETL) exercise where existing data is:

- Extracted from the old system
- Transformed to a format supported by the replacement system
- Loaded into the new system

### SERVICES DELIVERED:

- Content Analysis
- Content Classification
- Content Migration
- Data Quality
- Master Data Management
- Metadata Modeling

ETL success is dependent on the accuracy, completeness, and standardization of the source system's data. Unfortunately, data in long-lived enterprise systems often fails these standards due to a number of factors that can include:

### ISSUE

Find and repair incompatible metadata so information migrates seamlessly into new application

- Numerous administrators and data contributors (users) over time
- Multiple system ownership changes (e.g. owning department changes) with varying business needs and motivations
- Geographically dispersed operations
- Evolving business drivers and regulatory environments
- Acquisitions and mergers that contribute dissimilar data to the system
- Lack of consistent governance, configuration standards, and user training



These issues often result in a costly and time-consuming remediation process.



## METADATA: GARBAGE IN...GARBAGE OUT

This client was looking to replace their obsolete physical records management system (RMS), implemented in the early 2000's, which managed approximately 25M records. During the life of this system, multiple migrations took place, frequently without data validation. One recent migration involved a massive relocation of records from multiple locations to a centralized facility, and a substantial portion of the records' metadata was not correctly updated to reflect the move. As a result, the system contained incorrect or invalid metadata.

In preparing for the migration to the new RMS, the client identified an estimated 8.5M records they suspected had invalid or missing metadata that would prevent them from migrating. Issues included:

- Duplicate barcodes
- Missing or incorrect retention codes
- Missing review dates
- Invalid storage locations
- Ongoing litigation holds

Planning to remediate only one-third of the migrated records was unacceptable so the client allotted time and resources to find and fix these problem records.

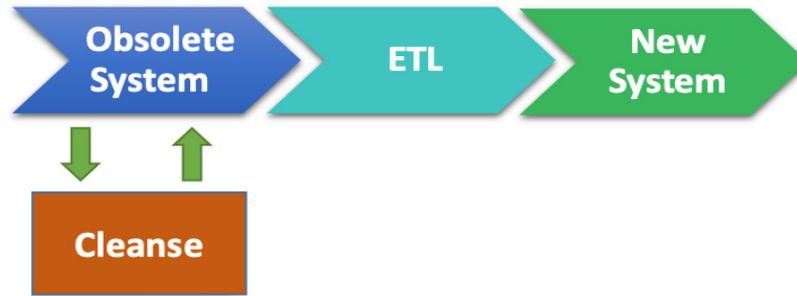
## CLEANSING THE DATA BEFORE MIGRATION

Rather than risk ETL failure or incur the expense of a prolonged post-ETL remediation, Access Sciences counseled the client to choose an alternative course of action, cleansing the data prior to ETL, ensuring a smooth transition into the replacement system.

---

## SOLUTIONS

- ✓ Migrated records to new RMS
  - ✓ Reduced migration costs by archiving 1.4M records
  - ✓ Eliminated cost of retired, obsolete system
  - ✓ Identified records eligible for destruction
  - ✓ Improved ability to retrieve content by correcting locations on 4.3M records
-



Our client, a global organization with operations in over 40 countries, trusted Access Sciences to cleanse their data prior to ETL. Our approach included:

- Parsing problem record sets according to issue, and assigning these to a dedicated team for resolution
- Utilizing tools to manage the large data sets needed to identify and resolve issues across the issue response teams.

---

## BENEFITS

- ☑ Data has been validated and standardized before implementing costly enterprise system
- ☑ Accurate data ensure staff can access records when needed

- Using a combination of resources and techniques to resolve issues that included information management best practices, data gathering, coffee, and elbow grease.

Our work processes revealed an additional 4M records requiring corrections, raising the overall total to 12.5M records. This represented a 47% increase in workload, which we were able to complete successfully on time and under budget.

---

## NEW SYSTEM STARTS WITH SOLID FOUNDATION

At the end of the data cleanse effort, the client's 25 million records had moved from 50% to 99.78% migration ready. With accurate source data in place, our client was able to confidently move forward with the migration to their new system.