

Real Estate Investment Trust (REIT) for Financial Transformation: an Iron Mountain Success Story

Objective of this Paper

Every organization will eventually need to undergo a transformation to compete in today's global economy. Using best of breed applications to fuel its growth into the next decade, Iron Mountain moved towards REIT enablement to gain global financial efficiencies as well as to redesign its internal business process. For this Information Security Company data quality and historical data conversion were key success factors for 99.99999% reconciliation and government audit before even beginning the growth strategy using Oracle Hyperion.

Intended Audiences:

(i) Individual contributor (ii) Project team member (iii) Project Manager

High Level Overview

Information quality problems and their impact are all around us: A customer does not receive an order because of incorrect shipping address, products are sold below cost because of wrong discount rates, a manufacturing line is stopped because the parts are not ordered due to incorrect inventory quantity. Information is not simply data, strings of numbers, list of addresses, or test results stored in a computer. Information is the product of business processes and is continuously used. Data quality problems may be caused by human, process, or system issues. Some normal business activities are indicative of data quality problems such as: correction activities, rework, reprocessing orders, handling orders and dealing with customer complaints. Both IT and business users must own the system and develop quality processes for data management.

Key Concepts

Information life cycle phases are Plan, Obtain, Store and Share, Maintain, Apply and Dispose. All phases of the Information life cycle have a cost. It is only when the resource is applied that the company receives value from it. Data quality is affected by activities in all of the phases of the life cycle. The major difference between information as a resource and other resources is

that Information is reusable; it is not consumed when used. If the information is wrong, it will be used again and again – with negative results.

Master Data: Master data describes the people, places and things that are involved in an organization's business e.g.: People (Customer, employees, vendors), places (locations, sales territories, organizations), and things (accounts, products, assets, document sets).

Reference Data: Reference data are sets of values or classification schemas that are referred to by systems, applications, data stores, processes and reports as well as by transactional and master records e.g.: Customer type in Customer Master data, Item type in Item master data.

Transactional Data: Transactional data describes an internal or external event or transaction that takes place as an organization conducts its business e.g.,: sales order, invoices, purchase orders, trips, deliveries, cash receipts, payments, inventory transactions etc.

Metadata: Metadata literally means "data about data". It shows all the characteristics of the tables and fields within them such as: Field name, Constraints, Data Type etc.

The Data Quality Improvement Cycle

Data can be improved with three steps within the cycle: assessment, awareness and action.

Assessment is looking at your actual data and comparing it to requirements and expectations, it **is key to Awareness**. **Awareness** is understanding the true state of your data, impact on the business and root causes, **which leads to Action**. **Action** is the prevention of future data quality problems in addition to correction of current data errors **which is verified by Periodic Assessment...so the cycle continues**.

The Process Steps

The 10 process steps followed as part of the Data Quality Improvement cycle are:

Assessment:

1. Define business need and approach
2. Analyze Information Environment
3. Assess Data Quality
4. Assess Business Impact

Awareness:

5. Identify Root Causes
6. Develop Improvement Plans

Action:

7. Prevent Future Data Errors
8. Correct Current Data Errors
9. Implement Controls

Assessment, Awareness and Action:

10. Communicate Actions and Results

While the ten process steps are appearing as linear processes, actually the process of data improvement is iterative – the project teams can return to previous steps to improve their work, the project team can choose the steps that the business needs and can repeat these steps as many times as needed to improve the data quality. Communication is critical and runs across all three life cycle phases.

1. Define Business Need and Approach:

Your project should only spend time on those issues where you expect to get results worth the time and money spent. In many cases you will find several issues or opportunities from which you will have to choose. You need to prioritize where to focus your efforts. If you have a favorite technique for prioritization, use it here.

Approach:

Consider issues or opportunities in the following areas such as:

- ≡ Lost revenue and missed opportunity – where revenue could increase if the data quality issue was addressed- for e.g.: increasing the sales with clean customer and contact information.
- ≡ Lost business – where your company once had a customer or vendor, but they chose not to do business with you because of some data quality issue.
- ≡ Unnecessary or excessive costs – due to excessive rework because the raw materials were not available due to incorrect inventory balances.
- ≡ Catastrophe – where poor data quality could lead to loss of life or some legal action.
- ≡ Increased risk – where data quality issues increase risk to your company – e.g.: when purchases of one customer are associated with duplicate customer master records causing the credit limit for that customer to be exceeded.
- ≡ Shared processes and shared data – where several transactions depend on the common data e.g.: Product data.

2. Analyze Information Environment:

Ensure that the requirements are relevant to the business issues, associated data, and data specifications necessary. Decide on the steps that you want to perform in the Data Quality Lifecycle process steps. Keep the scope manageable.

- ≡ Identify the business terms and data subject areas or groupings relevant to the business issue and associated requirements.
- ≡ Identify the systems/applications/databases where the data is used and the corresponding databases where the data are stored.
- ≡ Correlate the business terms or data subject areas with the specific fields that store the data in the database.
- ≡ Create a detailed data list for the data of interest
- ≡ Collect relevant data specifications for each of the fields of interest.
- ≡ Understand and document the relevant data model, entities, and relationships.
- ≡ If more than one data store is being assessed and compared, create a detailed data list for each one.
- ≡ Document additional information needed for the assessment.
- ≡ Document any potential effects on data quality or business impacts recognized at this time.

3. Assess Data Quality

Evaluate data quality for the data quality dimensions applicable to the issue. The assessment results provide a basis for future steps, such as identifying root causes and needed improvements and data corrections.

- ≡ Finalize your data capture and assessment plan
- ≡ Access or extract the data
- ≡ Profile the data
- ≡ Analyze the results
- ≡ Document results and recommended actions

Test, analysis and action are performed for number of records, completeness or fill rate, Nulls, List of unique values, Validity, Frequency distribution, Range of values, frequency, content, duplication, data type, size or length, patterns, precision, consistency, concurrency and timeliness, and business rules.

Duplication:

Checking for duplicates is the process of determining uniqueness. Automated tools can help you to perform this check the process is called data cleansing. “Matching” is a term generally used for linking the possible duplicate identifiers such as same customer, employee, supplier or product.

- ≡ True Match – Business rules as implemented within the tool have identified a match that is confirmed by business review
- ≡ Non-match – Business rules as implemented within the tool have identified unique records that are confirmed by business review
- ≡ False Negative – cases have been classified as unmatched but should have been matched– that is missed matches.
- ≡ False Positive – cases have been incorrectly classified as matches. They are actually nonmatches.

4. Assess Business Impact

Using a variety of techniques, determine the impact of poor-quality data on the business. This step provides input to establish the business case for improvement, to gain support for data quality and to determine appropriate investments.

- ≡ Collect examples or stories about the impact of poor data quality.
- ≡ Inventory the current and/or future uses of the data.
- ≡ Ask “why” five times to get to the real business impact.
- ≡ Rank the impact of missing and incorrect data on specific business processes.
- ≡ Illustrate the effects of poor-quality data.
- ≡ Compare potential benefits of investing in data quality with anticipated costs.

5. Identify Root Causes

Identify and prioritize the true causes of the data quality problems and develop specific recommendations for addressing them.

- ≡ Ask “why” five times to get to root cause.
- ≡ Identify location of the problem by tracking data through the information life cycle and determining root causes where the problem first appears.
- ≡ Identify, explore, and geographically display all possible causes of an issue by using a standard quality technique.

6. Develop Improvement Plans

Finalize specific recommendations for action. Develop and execute improvement plans based on recommendations.

- ≡ Gather results from each of the assessments and root cause activities.
- ≡ Develop and prioritize specific recommendations to address the issues found.
- ≡ Identify accountability and develop plans.
- ≡ Document and communicate results.

7. Prevent Future Data Errors

Implement solutions that address the root causes of the data quality problems.

- ≡ Ensure that improvement activities are focused on root causes.
- ≡ Implement appropriate improvement plans.
- ≡ Ensure the improvements on the data quality do not affect any other functionality.
- ≡ Project planning and effective communications.
- ≡ Develop and implement a data governance policy and process

8. Correct Current Data Errors

Implement steps to make appropriate data corrections.

- ≡ Identify the records to be updated or deleted.
- ≡ Decide how to make the changes without affecting existing functionalities.
- ≡ Determine who will make these changes. Manual or Automated tools.
- ≡ Make the changes. Document the changes. Communicate the results.

9. Implement Controls

Monitor and verify the improvements that were implemented. Maintain improved results by standardizing, documenting and continuously monitoring successful improvements.

- ≡ Use the data quality dimensions to control the data quality.
- ≡ Survey knowledgeable business users for issues in the data.
- ≡ Implement data governance, assign ownership and stewardship for each data element
- ≡ Ensure the Business has buy-in for whatever you do related to data quality controls.
- ≡ Measure the data quality improvements that have been implemented.
- ≡ Identify next potential area of improvements. Communicate the results.

10. Communicate Actions and Results

Document and communicate the results of quality tests, improvements made, and results of those improvements. Communication is so important that it is part of every step.

- ≡ Determine who needs to be included in your communication.
- ≡ Increase your communications skills.
- ≡ Agreement and consensus.

Other Techniques and Tools

Data Quality projects involve usage of both manual and automated tools. You need to capture the data for review (Data Extraction), check for possible issues with the data (Data Profiling/validation), Clean the data using cleansing algorithms (Data Cleansing), Manual review of the output of the Data cleansed, Manual Approval, update the data and return it to the system (Data Management).

- ≡ Data Extraction
- ≡ Data Profiling
- ≡ Data Validations
- ≡ Data Cleansing
- ≡ Data Management

These steps are iterative and can be utilized on a set time frequency.

Conclusion

From this paper you have learned several key concepts and the data quality improvement life cycles. The process steps within the data quality improvement life cycles can be used in any data quality project and the steps can be included in your project as appropriate. The concept of “Get Clean” and “Stay Clean” Data quality projects with the right automated tools and the right project team members is the key for your Project Success.