# Data Cleansing

A practical framework for identifying, remediating, and preventing master data inconsistencies in SAP landscapes

## Executive Summary

Enterprises running SAP ECC and SAP S/4HANA often face a repeatable pattern: inconsistent master data causes transaction failures, partner integration defects, reporting errors, and high remediation effort. Data cleansing is not a one-time migration activity—it must operate as an ongoing quality discipline. This white paper presents a structured approach to data cleansing that combines business-rule validation, data model integrity checks, best practices, automation-first remediation, and guided workflows for exceptions.

> *Guiding principle: run continuous audits ("data dialysis") so issues are detected early and corrected before they impact integrations and reporting.*

## Why Data Cleansing Fails in Practice

Most cleansing efforts fail because they focus only on fixing what is broken today, instead of building a system that prevents tomorrow's defects. Common failure modes include fragmented rule ownership, inconsistent validation across systems, manual remediation without traceability, and lack of governance for new fields introduced by industry frameworks or new data models such as Business Partners.

- Business-rule gaps: rules are tribal knowledge, not codified and versioned.
- Model drift: new fields and structures are introduced without harmonized validation.
- Manual remediation overload: stewards spend cycles identifying and rechecking issues.
- No continuous auditing: issues resurface because checks run only during projects.

## Core Concepts from the DMAG Approach

The DMAG approach frames cleansing as a combination of inconsistency identification and multi-mode resolution. Inconsistencies typically stem from three sources: (1) existing data violating proven business rules, (2) new data models such as Business Partners, and (3) new fields introduced by industry frameworks. Resolutions are similarly three-fold: automation for known issues, guided remediation for model/framework changes, and interactive workflows for steward-driven updates.

## A Reference Data Cleansing Lifecycle

| Phase | Objective | Typical Outputs |
|---|---|---|
| 1. Detect | Find inconsistencies using rules, model integrity checks, | Exception lists, validation |

| | | |
|---|---|---|
| | and best practices. | reports |
| 2. Diagnose | Classify issues by object, root cause, and ownership. | Issue taxonomy, priority scoring |
| 3. Remediate | Fix issues via automation where possible; guide users where not. | Automated updates, steward work queues |
| 4. Prevent | Embed validations upstream in processes and integrations. | Pre-save checks, reference data controls |
| 5. Monitor | Run audits on a schedule; track KPIs and drift. | Dashboards, trend reports, SLA alerts |

## Scope: What to Cleanse

Data cleansing should be scoped to the master data objects that drive the majority of downstream transactions and external distribution. Common objects include Materials, BOM, Recipes, Production Versions, Business Partners, General Ledger, Cost Centers, Profit Centers, and Employees. Start with the objects that have the highest transaction volume and integration blast radius (materials and business partners are typically first).

## Remediation Modes

A balanced program uses three remediation modes to combine speed and control:

| Mode | When to use | What to ensure |
|---|---|---|
| Automation (system-driven) | Auto-fix known issues where correct values can be derived deterministically. | Audit trail, validation, rollback plan |
| Guided remediation (steward-driven) | Export exceptions, allow stewards to correct values, then re-upload via controlled process. | Field help, validation, approvals |
| Interactive remediation (screen-driven) | Stewards correct records directly in a UI with validations and field help. | Role-based access, traceability |

## Case Study: Material Master Inconsistency Report

Material master data is critical for digital product information distribution and synchronization between business partners. A cleansing program flags missing or inconsistent values and produces inconsistency reports that can be routed to automation, guided remediation, or interactive steward workflows.

iv.   *Speed to Master Data Implementation by 50-60%%*

v.    *Reduces Master Data Stewarts FTEs – ON GOING - 60% FTE Savings*

## Product Information – A Case in Point

*Material Data is a critical component of Digital Product information distribution and for data synchronization between the business partners. There are number of considerations including Trade groups, Commodity Codes, Classification, GTINs, Customer Product Numbers, Attributes and Price characteristics. DMAG takes into account the above considerations and identifies all changes, provides a report of all inconsistencies including the new data elements , changed data elements, and Business Rule Inconsistencies. Business Rules are maintained in a custom table, so each client can define their business rules according to their Industry standards. Below is an example of Product Master inconsistency report. The inconsistency report shows '* **Business Rule inconsistencies'** *for each Product vs. Industry framework below in Figure(2).*



Figure 2 - Material Master Inconsistency Report

*Figure: Example material master inconsistency report (from the attached DMAG reference).*

*Figure: Example material master inconsistency report (from the DMAG reference).*

## Implementation Blueprint

| Step | What you do |
|---|---|
|  |  |

| | |
|---|---|
| 1) Identify & catalogue rules | Collect business rules by object; map each rule to fields, ownership, and severity. |
| 2) Build inconsistency detectors | Run rule checks and integrity constraints; output exceptions with context. |
| 3) Automate deterministic fixes | Auto-apply fixes where correct values are derivable (formatting, defaulting, mappings). |
| 4) Guided remediation workflow | Export exceptions with field help; ingest corrected files; validate; upload changes. |
| 5) Interactive remediation UI | Provide steward screens for complex issues; validations run at save time; approvals where needed. |
| 6) Continuous monitoring | Track KPIs, repeat defects, and rule effectiveness; tune rules over time. |

## Conclusion

Data cleansing becomes sustainable when it is operationalized: rules are governed, audits run continuously, and remediation balances automation with guided steward workflows. With this lifecycle, organizations maintain business-ready data, reduce manual effort, and improve partner integrations and analytics reliability.

### Contact

VisionSoft Global • 186 Princeton Hightstown Road, Princeton, NJ 08550 • info@visionsoft.global • +1 609 759 2967