Your Clean Core Is Only as Clean as Your Data

by **Uday Rao** August 27, 2026

Clean Core | S/4HANA | SAP Migration | Data Quality | Master Data | CVI | S/4HANA Migration | SAP HANA Licensing



Every S/4HANA program has a clean core roadmap. Fewer modifications, standard processes, cloud-ready extensions on BTP rather than code buried in the core. Far fewer programs fund the data workstream with the same seriousness, and that is the gap this paper is about. Your core can be immaculate on the process and code side and still run on duplicates, dead records and mismatched reference values from day one.

What does poor data quality actually cost?

Thomas C. Redman, writing in MIT Sloan Management Review, estimates the cost of bad data at 15 to 25 percent of revenue for most companies. The more useful half of his finding is that roughly two thirds of that cost can be identified and permanently eliminated once the root causes are addressed. This is not a permanent tax. It is a fixable, one-time investment with a payback.

Gartner puts the same problem in operational terms: poor data quality costs the average organisation 12.9 million dollars every year. In an SAP estate that shows up as delayed invoicing, mismatched vendor records, regional numbers that will not consolidate, and go-lives that stall in UAT because nobody looked at the data six months earlier.

Does CVI have to be completed before S/4HANA conversion?

Yes. Customer-Vendor Integration converts legacy Customer and Vendor master records into the Business Partner model that S/4HANA requires, and SAP's Software Update Manager will stop and fail the conversion if it is not complete. Every duplicate tax ID and every inconsistency between a Customer record and its matching Vendor record has to be reconciled before the technical conversion can proceed. SAP already agrees that data quality is not optional. It simply enforces it at the CVI checkpoint rather than trusting governance to catch it earlier.

Why does timing decide the cost?

The 1-10-100 rule, first articulated by Labovitz and Chang in 1992, holds that a data defect costs about one dollar to prevent at the source, ten dollars to correct once it is inside a process, and a hundred dollars to fix once it has reached production. Mapped onto an S/4HANA timeline: catching a duplicate vendor during ECC clean-up is the one dollar fix, catching it in integration testing is the ten dollar fix, and finding it after go-live inside three months of financial postings is the hundred dollar fix. By then it is a credibility problem rather than a data problem.

Download the full paper for the HANA licensing implications, the master data and reference data clean-up sequence, and what to put in front of your steering committee.