

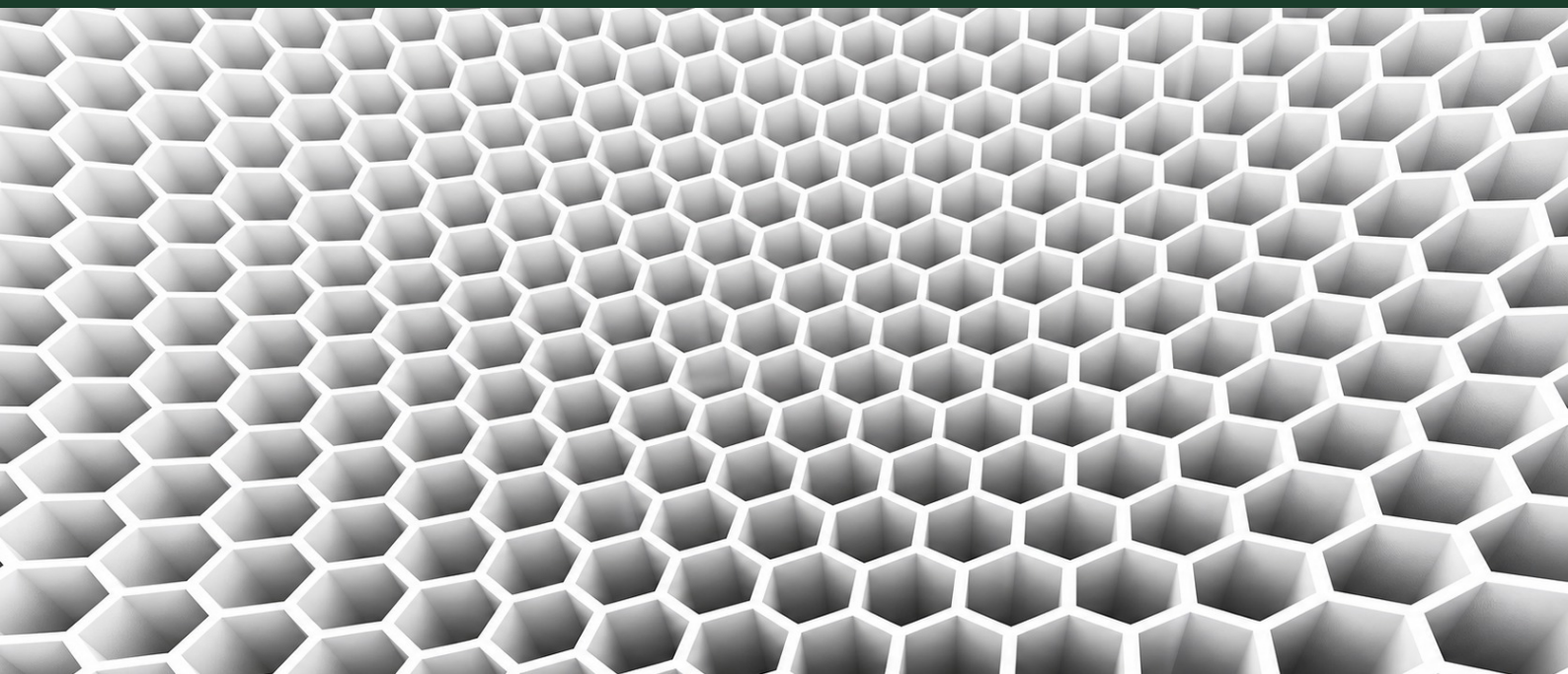
BEYOND THE CLEANSING SPRINT

Data Quality Resists Project Thinking

A Practitioner's Perspective on Why Data Quality Resists Project Thinking

Giovanni Leonardi

February 2019



Data Quality Resists Project Thinking

Programme · No. 182

Originally written February 2019 | Re-edited and re-rendered for the WhereBeyond Edition,
2025

“Data quality is not a problem to be solved; it is a condition to be managed, and the organisations that grasp this distinction are the ones whose transformation programmes actually stick.”

1. The Unsexy Truth About Transformation

There is no conference keynote on data quality. Nobody builds a career on it, nobody wins an award for it, and no consultancy leads with it in a pitch deck. And yet, if you were to ask me which single capability most reliably predicts whether a transformation programme will deliver lasting value, I would say data quality — without hesitation and without caveat.

The pattern I have observed across sectors — financial services, utilities, public sector, retail — is remarkably consistent. Organisations invest heavily in new platforms, new operating models, new customer journeys. They restructure teams, retrain staff, and redesign processes. And then, somewhere between the third and sixth month of operation, the new platform starts producing outputs that nobody trusts. The dashboards show numbers that contradict the old reports. The customer records have duplicates, gaps, and conflicts that the migration was supposed to resolve but merely relocated. The regulatory submissions require manual overrides that negate the efficiency gains the programme was built to deliver.

The root cause is almost always the same: data quality was treated as a workstream within the programme, not as a programme in its own right.

2. Why Project Thinking Fails Data Quality

The instinct to treat data quality as a project is understandable. It has a defined problem — the data is poor. It has identifiable deliverables — cleansed records, validated fields, reconciled sources. It appears to have a natural end state — clean data. Every element of the project management playbook seems to apply.

But data quality is not a state to be achieved; it is a condition to be maintained. The moment a project team declares victory and disbands, the entropy begins. New data enters through channels the project did not anticipate. Business rules change in ways that invalidate the cleansing logic. Staff turnover means the tacit knowledge of why

certain fields matter — and how they relate to one another — walks out of the building.

In my experience, the half-life of a data quality project's outcomes is between six and eighteen months. After that, the organisation is back where it started, often with the additional complication of having lost the institutional memory of what was done and why.

3. The Programme Model

The organisations that get this right — and they are fewer than the industry would like to admit — treat data quality as a standing programme with its own governance, its own budget cycle, and its own career paths. This is not a theoretical ideal; it is a structural necessity that follows from the nature of the problem.

A data quality programme differs from a data quality project in several fundamental respects:

- **Permanence over closure.** The programme does not have an end date. It has review cycles, maturity milestones, and evolving objectives, but it operates on the assumption that data quality is a perpetual concern.
- **Ownership over execution.** The programme defines and enforces data ownership — not in the abstract sense of a RACI chart filed in a SharePoint library, but in the operational sense of named individuals who are accountable for the quality of specific data domains and who have the authority to reject data that fails to meet defined standards.
- **Prevention over remediation.** The programme invests disproportionately in input controls — validation at the point of entry, integration standards between systems, and automated quality rules that catch degradation before it propagates — rather than in periodic cleansing exercises that treat symptoms.
- **Measurement over anecdote.** The programme maintains quality metrics that are reported with the same rigour and regularity as financial metrics. Data quality scores appear in board packs, in operational reviews, and in programme health dashboards — not as a curiosity, but as a leading indicator of operational risk.

4. What GDPR Revealed

The General Data Protection Regulation, which came into force in May 2018, has done more to expose the consequences of poor data quality than any technology initiative or consulting engagement. Organisations that could not reliably identify what personal data they held, where it was stored, or how it flowed between systems found themselves unable to meet the most basic requirements of the regulation — not because they lacked the will, but because the data itself was not in a condition to be governed.

This is the connection that many compliance programmes have missed. GDPR is not primarily a legal problem or a technology problem; it is a data quality problem. The organisations that had invested in data quality as a programme — that could trace data

lineage, validate accuracy, and demonstrate provenance — found GDPR compliance to be a manageable exercise. Those that had not found themselves running emergency remediation projects that consumed budgets intended for strategic investment.

The lesson is straightforward but uncomfortable: regulatory compliance is a consequence of good data management, not a driver of it. Organisations that only invest in data quality when a regulator forces their hand will always be reactive, always behind, and always spending more than those that treat it as a core operational capability.

5. The Governance Gap

The most common failure pattern I observe is what I think of as the governance gap — the space between the data quality standards that exist on paper and the data quality practices that exist in operation.

Most large organisations have a data governance framework. They have policies, standards, and often a Chief Data Officer with a mandate to improve data management. What they frequently lack is the operational machinery to make these frameworks effective at the point where data is created, modified, and consumed.

Data quality is not a problem to be solved; it is a condition to be managed, and the organisations that grasp this distinction are the ones whose transformation programmes actually stick.

The governance gap manifests in predictable ways. Data quality rules are defined centrally but not enforced locally. Data stewards are appointed but given no time allocation, no tools, and no authority. Quality dashboards are built but not acted upon — they become wallpaper, observed but not interpreted.

Closing this gap requires something that governance frameworks alone cannot provide: a programme structure that connects policy to practice through standing teams, funded capacity, and clear escalation paths. It requires, in short, treating data quality as programme management treats any other ongoing operational concern — with dedicated resources, defined accountabilities, and regular performance review.

6. The Career Problem

There is a structural reason why data quality struggles for attention, and it is worth naming directly: there is no career in it. The incentive structures of most organisations reward the visible and the novel — the new platform, the digital channel, the innovation lab. Nobody gets promoted for maintaining data quality. Nobody gets headhunted for a track record in data stewardship.

This creates a vicious cycle. The most capable programme managers and analysts gravitate towards more glamorous work. Data quality is staffed with whoever is available, or worse, treated as a part-time responsibility layered onto roles that already have a full

mandate. The work suffers, the outcomes disappoint, and the next time investment is requested, the case is harder to make because the last initiative did not deliver.

Breaking this cycle requires a deliberate choice by senior leadership to signal — through budget allocation, through recognition, and through career development — that data quality is a first-class discipline. In the organisations where I have seen this work, it has always started with a senior sponsor who understood, from direct experience, the cost of getting it wrong.

7. What Good Looks Like

The organisations that run data quality as a genuine programme share several characteristics that are worth noting, not because they are surprising, but because they are so rarely implemented together:

- They have a standing data quality team with its own budget line, not a virtual team assembled from other functions.
- They measure data quality continuously, not periodically, using automated profiling that runs against live operational data.
- They treat data quality failures as operational incidents, with the same severity classification, root cause analysis, and corrective action processes applied to system outages.
- They connect data quality metrics to business outcomes — customer complaints, regulatory findings, operational errors — so that the case for investment is grounded in consequence, not abstraction.
- They invest in data literacy across the organisation, so that the people who create and consume data understand why quality matters and what their role is in maintaining it.

None of this is glamorous. None of it will feature in a case study or a conference presentation. But it is the foundation upon which every other transformation ambition rests, and the organisations that recognise this — that treat the boring work as the essential work — are the ones whose transformations endure.