

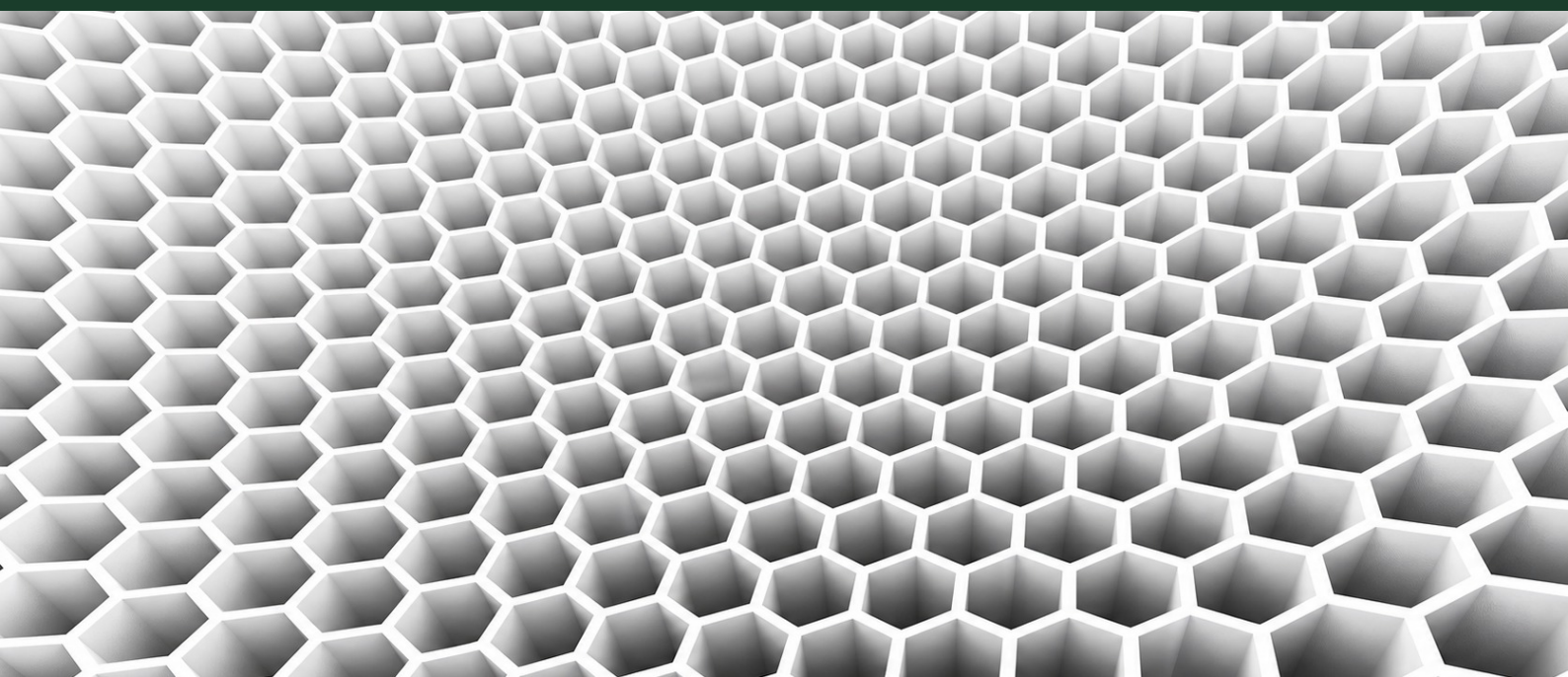
BEYOND THE CLEAN-UP PROJECT

The Most Boring Transformation That Matters Most

A Practitioner's Perspective on Why Data Quality Is a Programme, Not a Project

Giovanni Leonardi

February 2019



The Most Boring Transformation That Matters Most

Programme · No. 181

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

“You cannot photograph the absence of a problem.”

1. The Transformation Nobody Wants to Lead

Every organisation I observe agrees, in principle, that its data is a mess. The agreement is genuine and it is nearly universal. Ask an executive whether the figures in their reports can be trusted end to end, and after a short pause you tend to get an honest answer — and the honest answer is usually no. Yet data quality remains the transformation that nobody wants to lead, that never quite gets funded properly, and that reliably reappears eighteen months later as though it had never been addressed at all.

I want to be blunt about why, because after enough cycles the pattern is no longer a mystery. Data quality does not fail because it is technically difficult. The techniques for profiling, cleansing, matching and monitoring data have been well understood for a long time. It fails because organisations persist in treating a permanent operating discipline as a temporary project — and no amount of tooling rescues a permanent problem that has been handed a finish line.

1.1 The project delusion

The word that does the damage is project. A project has a start, an end, a budget that closes, and a team that disbands. Frame data quality that way and you have already lost, because the day after the project celebrates its completion, the organisation carries on creating data exactly as it did before. New records are keyed with the same shortcuts. The same fields are left blank. The same three systems continue to disagree about the same customer. Quality decays from the moment the cleanse finishes, because nothing about how the data was produced has changed.

“A cleanse fixes the data you have. Only a programme fixes the data you are about to create.”

This is the distinction that gets lost. Cleansing is a treatment of symptoms — valuable, necessary, and utterly temporary. The disease is in the processes, incentives and ownership that produce poor data in the first place, and those are not things a

time-boxed project is equipped to touch. A programme can. A programme persists, holds a mandate over how data is created and maintained across its whole lifecycle, and treats quality as an operating standard rather than a milestone. This is why the same organisations run the same remediation every eighteen months and call it new each time: they keep curing the symptom and leaving the cause untouched, then express surprise when the symptom returns.

- A project asks: how many records can we cleanse before the deadline?
- A programme asks: why did those records become wrong, and what has to change so the next million are born correct?

1.2 Why the honest account is uncomfortable

Here is the part the textbooks leave out. The reason data quality is chronically under-led is not ignorance. Everyone knows it matters; it has become fashionable to say that data is the organisation's most valuable asset, and people say it with feeling. The reason is that data quality is unrewarding to own. It has no launch moment, no glamorous demo, no ribbon to cut. Its success is invisible — a report that was simply right, a duplicate that never existed, a regulatory return that filed without incident. You cannot photograph the absence of a problem. Careers are not made on transformations whose highest achievement is that nothing went wrong.

There is a quiet hypocrisy in this. The same leaders who call data their most valuable asset decline to fund its upkeep the way they would fund the upkeep of any other asset they genuinely valued. An asset nobody is willing to maintain is not being treated as an asset at all. It is being treated as a slogan.

So the work becomes the responsibility that is always delegated downward — handed to IT because the data lives in systems, as though the systems were the ones filling the fields in incorrectly. This is perhaps the single most common structural error I see. Data quality is placed with the function that has the least influence over the behaviours that create the data. IT can build the pipes and the profiling dashboards, but it cannot make a sales team enter a complete address, or persuade three business units to agree on what a customer even is. Ownership sits in the wrong place, and so accountability quietly evaporates.

Data quality is the only transformation routinely assigned to the function with the least control over its outcome. We give it to the people who store the data rather than the people who create it, and then we are surprised when nothing improves.

1.3 The regulator has changed the conversation — for better and worse

For years, the business case for data quality was a genuinely hard sell, because the cost of poor data is diffuse and the benefit of good data is hard to attribute to any single decision. That changed over the past year. The arrival of the new data protection regime has done what a decade of internal advocacy could not: it has given data quality a sponsor at board level. The accuracy principle is now a legal obligation rather than a nice-to-have. The right to have inaccurate personal data corrected is enforceable. Suddenly the mess has a price, and the price has the board's attention.

I welcome the sponsorship. But I would be dishonest if I did not name the risk that comes with it, because it is already visible. When fear of enforcement becomes the motive, organisations build the wrong programme — one scoped to personal data alone, aimed at demonstrable compliance rather than genuine fitness for use, and destined to close the moment the audit is passed. That is the project delusion again, wearing a compliance badge. It will cleanse what the regulation can see and leave everything else exactly as broken as it was.

Worse, it manufactures false confidence. A board that has passed its compliance audit believes its data house is in order, when all it has really established is that the subset of data the regulator cares about has been tidied for inspection. The operational data that runs the business — the pricing, the inventory, the reconciliations — may be as unreliable as ever. Compliance-shaped assurance is the most dangerous kind, because it looks exactly like the real thing from the boardroom.

“Compliance will fund the programme. It will not, on its own, design a good one.”

The organisations getting this right are using the regulatory pressure as the reason to start while refusing to let it define the shape of what they build. They treat the accuracy obligation as the thin end of a much larger wedge: if we must guarantee the quality of personal data, we may as well fix the operating model that degrades all of it. The regulation becomes the funding argument, not the specification.

1.4 What a programme actually looks like

If data quality is a programme rather than a project, then it has the properties of a programme, and they are worth stating plainly because they are so often absent:

- 1. It has a permanent owner with business authority — someone senior enough to change how data is created, not merely how it is stored. This is why the slow emergence of a genuine data leadership role at executive level matters more than any tool an organisation could buy.
- 2. It is measured by decisions improved, not records cleansed. The only metric that means anything is whether the people who rely on the data now trust it enough to act on it without a manual check first.
- 3. It owns the whole lifecycle, and the point of creation above all. Quality designed in at the moment of capture is worth an order of magnitude more than quality inspected in afterwards, because prevention scales and inspection does not.
- 4. It never finishes. It moves through remediation, then prevention, then monitoring, but it does not disband, because the organisation never stops producing data.

The contrast with how these initiatives are usually run is not subtle:

How data quality is usually run	How it has to be run
Funded as a project with an end date	Funded as a standing operating capability

How data quality is usually run	How it has to be run
Owned by IT, because data lives in systems	Owned by the business, because behaviour creates data
Measured by records cleansed	Measured by decisions the data can be trusted to support
Scoped to whatever the audit can see	Scoped to the operating model that produces the data

1.5 The unglamorous conclusion

I called this the most boring transformation that matters most, and I meant both halves sincerely. It is boring. It produces no dramatic before-and-after, and its champion will never be celebrated the way the launch of a new platform is celebrated. It also quietly underwrites everything else. Every analytics ambition, every automation, every confident board decision rests on data somebody assumed was correct. When the foundation is wrong, everything built on it inherits the flaw, and nobody can quite say why the results feel untrustworthy.

The regulator has, almost by accident, handed us the best opportunity in years to fix this properly. Whether organisations take it comes down to a single, unglamorous choice: whether they fund data quality as one more project that will be quietly forgotten, or finally accept it for what it has always been – a permanent discipline that has no end, wants no ribbon, and quietly determines whether anything else they attempt can be believed.