DIRTY DATA EXPERT

Structure Before Intelligence™

THE QUESTIONER

Executive Brief

Gary Segal
www.dirtydataexpert.com
gary@dirtydataexpert.com

The employee your AI is waiting for

LeniG's professional background is in fashion design. To support her design work, she developed her own approach to trend forecasting — Google Trends, five years of data at a time, and her own eyes.

Working on footwear trends for Autumn/Winter 2027–28, she was analysing five years of search data when one graph stopped her. For almost the entire period, activity was extremely low. Then, in a specific period in 2026, it spiked dramatically — and collapsed again. She didn't build the forecast anyway. She investigated, and found independent reporting of similar spikes across hundreds of unrelated search terms. Something was wrong with the data itself.

Her words: "Sometimes the most important discovery isn't what the data tells us, but what makes us question the data in the first place."

In her other work, entering orders, she selects the warehouse in two fields. The second selection sometimes changes unexpectedly. The system accepts the order without warning. She corrects it — and then notifies the other warehouse, passes the order reference, and keeps the issue open until the duplicate entry is confirmed deleted. Her words again: "Correcting an error doesn't necessarily mean correcting all its consequences."

LeniG is the rarest person in your business. I call her the Questioner. And in forty years of working with data, I have learnt that almost nobody else does what she does.

Square one

People are wary of computers. They always have been.

In 1980, when the 5.25-inch floppy disk was the main storage medium, an accountant told me he had dropped a disk and asked whether the data was now corrupted. He imagined the data as a physical thing that could fall out and break.

Later I implemented business intelligence tools in finance and retail, and a software suite called Scala, built on OLAP. I promoted the drill-down concept — go past the summary, see what the data is actually saying. It was not used. The complaint was time. I don't believe it was time. Self-taught users only work inside their own environment, the small routine they were shown. They are scared to delve into the data and see what it tells them. A computer science graduate is not afraid of experimentation, because he holds a mental model of how the machine works. The rest hold a superstition.

Now everybody is using AI because it is the in thing. The wariness didn't change — only the tool did. Most users never leave square one: they accept the output the way they accepted the summary report, the way they feared the floppy disk. The few who move to square two — who question the input before trusting the output — are the only ones giving AI the justice it needs.

The cost nobody counts

Executives will tell me the Questioner is slow. She keeps issues open. She notifies, confirms, double-checks. Multiply her and throughput drops.

Count the other side. If LeniG were incentivised to draw management's attention to the software bug behind the warehouse error, the company would save the time and cost of reconciling discrepancies

forever. A discrepancy in the data is dirty data, and dirty data is reconciled every month, every year-end, every audit — at full price, repeatedly. The Questioner's minutes are the cheapest line in that chain.

And where does the dirty data come from? Management, migrating software platforms, looks for cost-saving measures, and shortcuts are taken. I was at an electronics firm that allowed the posting of accounting transactions spanning across financial years. Recipe for disaster. The disaster wasn't created by a slow employee asking questions. It was created in a boardroom, and paid for in reconciliation for years.

AI will take the blame

Understand what AI changes and what it doesn't. It changes nothing about the cost — the cost of unquestioned data was always there. It has a secondary impact: the AI intervention, and the information that results from it, will be flawed — and AI will be blamed.

This is not a prediction. It is already the record. A Federal Reserve Bank of Atlanta working paper surveyed almost 6,000 firms across the US, UK, Germany and Australia: around 70% actively use AI, yet more than 80% report that it has had no impact on their productivity or their employment over the past three years. Companies that rushed in are pulling back — restricting the tools, bringing back the experienced workers they let go. And the conclusion being drawn, in boardroom after boardroom, is that AI doesn't work.

Not the data. Not the shortcuts. The model. The company will conclude that AI doesn't work, abandon the tool, and leave the foundation exactly as it was — ready to fail the next intervention in exactly the same way. If the software platform is not bulletproof, you are chasing a moving target. That is a failed AI intervention before it starts.

Structure before intelligence

So build the structure that protects the Questioner instead of crushing her.

The Questioner reports to the CFO — the executive who owns the numbers and personally carries the cost when the data is wrong. Her platform integrity sits with the CIO, who holds the programmer — or the person who signs off the code — responsible. Not goodwill. A reporting line, and a name on a sign-off.

When AI fails, it is the data that is wrong, for starters. Somebody must be accountable for that data before it reaches the model, and somebody in the building must be employed to question it.

Hire the attribute

You do not find Questioners by testing for finished skill. You look for people who have attributes that can be molded into the worker you want.

When I interview a person for a sales role, I don't want to know how intelligent she is — she will try to impose her knowledge. I want someone with a smile in their voice. Training and guidance can come later. The attribute cannot.

Your current filters do the opposite. Degree in the field. Python. BI tools. LeniG fails all three — and she is the sharpest data instinct in her building. Companies are screening out the exact persona their AI depends on, then wondering why the AI fails and blaming the model.

Hire the questioner. Mold her. Give her a line to the CFO and a reason to speak. The model was never your problem.

Structure before intelligence — and people before structure.

Sources: Federal Reserve Bank of Atlanta, Working Paper 2026-3, "Firm Data on AI" (survey of almost 6,000 firms, US, UK, Germany, Australia); company pullback reporting, The Guardian, 13 October 2026.

Gary Segal is the author of Dirty Data and Switching ERP for Embedded AI. He writes the weekly Structure Before Intelligence™ executive brief at dirtydataexpert.com.