ATELIER VALUE

WHITEPAPER · 5 MIN READ

Why AI Pilots Die on the Shop Floor

Rafael Herrera, Founder & Managing Director · November 2026

> **Core insight:** Supply chain AI rarely fails in the demo. It fails in the handover, when a model built on clean data meets a plant that runs on exceptions, conflicting targets and unclear decision rights. Scaling it is an operating-model problem, not a model-selection problem.

1. The pilot that looked like a success

The steering committee has seen the slide. A forecasting or scheduling model, trained on a year of history, beat the planners' baseline on the test set. The data science team is proud. Everyone agrees to scale it next quarter.

Six months later, the model still runs. Its output lands in a shared folder every Monday. The planners open it, compare it with their own spreadsheet, and keep the spreadsheet. No order, schedule or shipment has changed because of the model. The pilot did not fail. It simply never arrived.

This is the quiet way supply chain AI dies: after the applause, somewhere between the model and the decision it was meant to make.

2. Why the plant is a different country

A pilot is designed to succeed. It runs on a curated set of SKUs, cleaned data, a friendly planner and a deadline measured in weeks. The plant and the planning desk run on something else:

- **Exceptions, not averages.** A line goes down, a supplier ships short, a key account moves its promotion. The pilot was scored on the normal week. The plant lives in the abnormal one.
- **Master data that decays.** Lead times, pack sizes, minimum batches and routings drift every month. A model is only as current as the last person who maintained them.
- **Conflicting incentives.** Planning is measured on service, manufacturing on efficiency, procurement on price, finance on inventory. A proposal that helps one target hurts another, so someone overrides it.
- **Tribal knowledge.** The best planner knows which line runs slow after a changeover and which customer always over-orders before a holiday. None of that is in the training data.
- **Unclear decision rights.** Nobody has written down who may accept, reject or change the model's proposal, or who carries the consequence when it is wrong.

> *“A model can be accurate and still change nothing. Value appears only when an order, a schedule or a shipment is different on Monday morning.”*
>
> RAFAEL HERRERA

3. What has to be true for a decision to run

Pilots are scored on model accuracy. Plants are scored on service, inventory and cash. That gap in scorecards is where the stall begins. Forecast accuracy can improve while service stays flat, because the forecast was never connected to the replenishment rule, the production plan or the purchase order.

So start with a different question. Not “which model?” but “which decision, owned by whom, executed where?”

DIMENSION	PILOT WORLD	PLANT WORLD
Scope	Curated SKUs, one site	Full portfolio, every exception
Data	Cleaned once for the test	Maintained weekly by named owners
Success metric	Accuracy on a test set	Service, inventory and cash
Decision owner	The project team	A named role with written authority
Exceptions	Excluded from the sample	Routed, prioritized and resolved
Integration	A file or a dashboard	Written back to ERP or MES
Overrides	Informal and invisible	Logged, reasoned and reviewed
Exit criterion	A positive steering committee	The decision runs without a meeting

The override log matters more than it looks. Every time a planner rejects the model, there is either a gap in the model or a gap in the planner’s information. Reviewed weekly, overrides become the fastest way to improve the model and to earn the planners’ trust. Ignored, they become the reason it is switched off.

4. The mandate

Start with one decision, not a platform. Choose one that is frequent, bounded and measurable: a replenishment proposal for one product family, the weekly allocation of constrained supply, the production sequence on one line. Write down who owns it, what the model may do alone, what it must escalate, and how success will be measured in service, inventory or cash. Connect the output to the system where the decision is executed. Then run it alongside the planners until the overrides say it is ready.

It needs people who have run planning and operations at scale, who can sit with the scheduler at six in the morning and understand why the recommendation was ignored. An atelier is a craftsman’s workshop: bespoke work, made with precision and finished by the people who designed it. We design the strategy, then stay on the front line until it runs.

The first closed-loop decision is the hardest. The second is faster, because the data owners, decision rights and override routine already exist. That is how a pilot becomes an operating capability. In the planning transformation I led at a global consumer goods multinational, that is how more than 40% of tactical planning decisions came to run without manual intervention.

Is your next AI investment another successful pilot, or the first decision your plant actually runs?

Before you fund the next pilot, a short, senior scan of your planning and plant decisions shows where AI can close its first loop, and what has to be true before it does. Connect directly with our team. Or start with the three-minute Supply Chain & AI Readiness Assessment.

Contact: ateliervalue.com/contact · Assessment: ateliervalue.com/readiness
Online version: ateliervalue.com/insights/why-ai-pilots-die-on-the-shop-floor