47 REAL PROJECTS ANALYZED

The AI Implementation Autopsy Report

What Actually Worked (And Why Most Failed) — The Post-Mortem Nobody Had the Courage to Write

JJ

JJ Shay
AI Strategy & M&A Executive

SLIDE 02

The Number Nobody Wants to Publish

73%	$4.6B	18
AI PROJECTS FAIL TO REACH PRODUCTION	WASTED ON FAILED AI INITIATIVES (2023)	AVERAGE MONTHS BEFORE FAILURE ADMITTED

> *Everyone publishes success stories. This guide is different — it's the honest failure analysis that vendors won't show you.*

SLIDE 03

Why This Report Exists

The Data

- **47** Real implementations analyzed across industries
- **89** Post-mortem interviews with project leads
- **$127M** Total investment represented in study

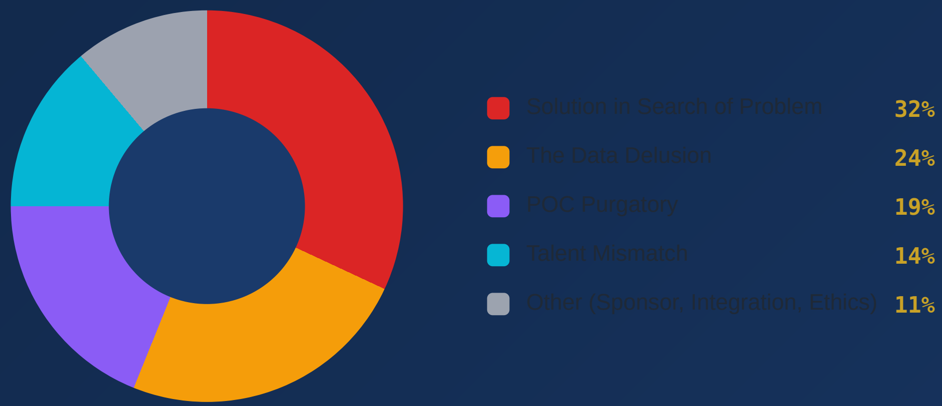
The Promise

By the end of this guide, you'll know:

- ✓ The 7 patterns that kill AI projects
- ✓ Warning signs visible in month 1
- ✓ How to run a pre-mortem before you start
- ✓ Kill criteria that save millions

SLIDE 04

The 7 Failure Patterns



> *56% of failures happen before a single line of ML code is written.*

SLIDE 05 — PATTERN #1

Solution in Search of a Problem

Starting with "we need AI" instead of "we need to solve X"

- ⚠ **"Our competitors are using AI"** — Fear-driven initiatives without clear business case
- ⚠ **"The CEO read an article"** — Top-down mandates without problem validation
- ⚠ **"We have this data, let's do something with it"** — Asset-first thinking vs. outcome-first

The Fix

Before any AI project, answer: "What specific decision will be made better, faster, or cheaper?"

SLIDE 06 — PATTERN #2

The **Data Delusion**

Assuming the data exists and is clean when it isn't

67%
PROJECTS DISCOVER DATA ISSUES AFTER KICKOFF

3.2x
AVERAGE DATA PREP TIME VS. ESTIMATE

- ⚠ **"We have 10 years of data"** — Yes, but in 47 different formats across 12 systems
- ⚠ **"Our data is in a data warehouse"** — Stored ≠ labeled ≠ ML-ready

SLIDE 07 — PATTERN #3

POC Purgatory

Successful pilots that never scale to production

MONTH 3

"The POC is working great! 94% accuracy in the lab."

MONTH 6

"We're working on productionization. Just need a few more integrations."

MONTH 12

"Still in POC phase. Exploring alternative approaches."

MONTH 18

"Project deprioritized due to changing business needs."

> *POCs are designed to prove concepts. They're not designed to prove scalability, maintainability, or business value.*

SLIDE 08 — PATTERNS #4-5

Talent & Sponsorship

Pattern #4: Talent Mismatch (14%)

- ⚠ Data scientists hired, but needed ML engineers
- ⚠ PhD researchers, but needed production builders
- ⚠ Junior team on senior-complexity problem

Pattern #5: Sponsor Disappearance (8%)

- ⚠ Executive champion leaves mid-project
- ⚠ Reorg shifts priorities
- ⚠ Budget owner changes strategy

SLIDE 09 — AUTOPSY #1

The $4M Chatbot Nobody Used

FORTUNE 500 COMPANY

Customer Service AI

Investment	$4.2M
Timeline	18 months
Expected ROI	$12M/year
Actual ROI	-$4.2M

The Setup

A Fortune 500 company invested in an AI chatbot to deflect 40% of customer service calls.

Team: Top-tier vendor + internal IT + Big 4 consulting

Technology: State-of-the-art NLP, custom training on 5 years of call transcripts

SLIDE 10 — AUTOPSY #1

What Went Wrong

MONTH 3

Technical POC succeeds. 89% intent recognition accuracy.

MONTH 8

Integration complete. Bot deployed to production.

MONTH 10

Usage data shows 78% of users abandon bot within 30 seconds.

MONTH 14

First user testing conducted. "Customers hate it."

MONTH 18

Project quietly discontinued. Team reassigned.

> *We built exactly what we specified. We just specified the wrong thing.*
>
> — Project Lead, Exit Interview

SLIDE 11 — AUTOPSY #1

Lessons Learned

- ✓ **User testing from Week 2, not Month 14.** Real users would have identified the UX problem immediately.
- ✓ **Define success as user adoption, not technical deployment.** The model worked. The product didn't.
- ✓ **Establish kill criteria upfront.** "If usage drops below X by Month Y, we pivot or kill."

The Core Mistake

They optimized for technical metrics (intent accuracy) instead of business metrics (call deflection). **A 99% accurate bot that nobody uses has 0% business value.**

SLIDE 12 — AUTOPSY #2

The Predictive Model That Backfired

MID-MARKET MANUFACTURER

Demand Forecasting ML

Investment	$1.8M
Goal	-23% inventory costs
Actual Result	$8M in losses

What Happened

The model recommended aggressive inventory cuts based on historical demand patterns.

Problem: Training data didn't include COVID-era demand volatility.

Result: Stockouts during demand surge → $8M in lost sales, customer churn, emergency air freight.

SLIDE 13 — AUTOPSY #2

The Missing Question

Nobody Asked:

"What's the cost of the model being wrong?"

- ✓ **Asymmetric risk analysis before deployment.** Understocking cost 4x more than overstocking.
- ✓ **Human-in-the-loop for high-stakes decisions.** Model should recommend, not decide.
- ✓ **Test for "black swan" scenarios.** What happens when the model encounters data it's never seen?

> *The model was technically excellent. It just couldn't handle reality.*

SLIDE 14 — AUTOPSY #3

The AI-First Startup That Almost Died

The Vision

- **$50M** Series B raised on "AI-native" positioning
- **$3M** Monthly burn, mostly AI infrastructure
- **18** Months of proprietary AI development

The Crisis

ChatGPT launched.

80% of their "proprietary AI" became table stakes overnight.

8 months of runway left. No clear pivot.

> *We were building the car while Tesla was building the assembly line.*
>
> — Co-Founder

SLIDE 15 — AUTOPSY #3

The Save

PIVOT STRATEGY

Workflows on LLMs

Old Model	**Build proprietary AI**
New Model	**Orchestrate foundation models**
Outcome	**Survived (60% workforce cut)**

The Lesson

- ✓ **Build ON foundation models, not AROUND them**
- ✓ **Your moat is domain expertise + workflows, not raw AI**
- ✓ **Monitor foundation model releases as strategic threats**

The Resurrection Playbook

12 of 47 "failed" projects were salvaged and became successes.

The 4-Step Recovery Framework

1. **Acknowledge Without Blame** →
2. **Salvage Assessment** →
3. **Problem Redefinition** →
4. **Rapid Retest**

> *Most AI failures are positioning failures, not technology failures. The tech often works — it's just solving the wrong problem.*

SLIDE 17

The Pre-Mortem Framework

Run this exercise before ANY AI project kickoff.

The Exercise

"Imagine it's 18 months from now and this project has failed spectacularly. Why?"

Ask Your Team:

1. What data assumption could be wrong?
2. Who might leave that would kill this?
3. What market change could make this irrelevant?

Ask Leadership:

4. What budget cut would kill this?
5. What would make us deprioritize this?
6. What's our actual risk tolerance?

SLIDE 18

AI Project Health Dashboard

8 leading indicators to track monthly

📊 DATA QUALITY SCORE	👥 USER ENGAGEMENT	⚡ MODEL PERFORMANCE	💰 BURN RATE VS. PLAN
🎯 BUSINESS KPI IMPACT	🔄 SCOPE CREEP INDEX	👔 SPONSOR ENGAGEMENT	🚦 INTEGRATION PROGRESS

If 3+ indicators are red for 2+ consecutive months → escalate immediately

SLIDE 19

The Kill Criteria Decision Tree

Define these BEFORE the project starts — not when you're $2M in.

THE TAKEAWAY

The Best AI Teams

Fail Fast.

The worst fail slow and expensive.

"

Define your kill criteria in the first project meeting, not the last.

Want the full framework + templates?

DM me "AUTOPSY" on LinkedIn →

JJ
JJ Shay
linkedin.com/in/jjshay