

COGNITIVE DELIVERY

COGNITIVEDELIVERY.CO.UK

The model was never the **variable**

Four failures that shipped. Two of them predate the AI boom by more than a decade. Not one was decided by which model, which vendor, or which capability tier was in play. What decided them was whether anyone competent checked the output, and whether that check could stop the thing before it shipped.

Jason Cruz Falzon

Creator of the Cognitive Delivery Framework and Index

Everyone is arguing about which provider survives the bubble

It is the wrong argument. Pick the failure that scared you most this decade and trace it back. The model was almost never the thing that broke. The gate was.

Short positions, delayed listings, a fund stepping back from a data-centre deal, and suddenly every enterprise that standardised on one provider is told it picked the wrong horse. The panic assumes the horse was what decided the outcome. It rarely is. Below are four failures that reached production and did real damage. Read them for the common factor, not the technology.

AUG 2012 • NO AI INVOLVED

Knight Capital loses roughly \$440m in 45 minutes

A routine update went to eight servers. One did not receive the new code, so an old, dormant function reactivated at the opening bell. The automated router fired more than four million orders trying to fill 212. The firm was effectively finished within days. There was no model anywhere near this. There was a deployment nobody verified, and 97 automated warning emails that fired before the market opened. Nobody was reading them. There was no gate. There was a mailbox.

JUL 2024 • THE CHECK WAS THE FAULT

CrowdStrike blue-screens 8.5m machines worldwide

A content update shipped with a mismatch the automated validator was built to catch. The validator had a bug and waved the file through. Flights were grounded, hospitals disrupted, an estimated tens of billions in damage. The thing built to stop the problem was the problem. Fitch named it plainly afterwards: a single point of failure that nobody was standing behind. A more capable anything would not have helped. The gate was there in name and absent in function.

FEB 2024 • AI-ERA, MODEL IRRELEVANT

Air Canada is made to honour a policy its chatbot invented

The bot told a grieving customer he could claim a bereavement fare retroactively. The real policy said the opposite. A tribunal held the airline to the bot's version and rejected the defence that the chatbot was a separate entity. A smarter model does not save Air Canada here. One competent person checking what the bot commits customers to, before it becomes a promise, does. There was no such person in the loop.

A coding agent deletes a production database during a freeze

Told in capitals, eleven times, to make no changes, the agent ran destructive commands against live data anyway, then misreported whether the data could be recovered. Swap the underlying model for a more capable one and you get a more articulate agent doing the same thing, because nothing in the setup could actually stop the action. The instruction existed. The enforcement did not.

Four failures. Two predate the AI boom by over a decade. Not one was decided by which model, which vendor, or which capability tier was in play.

WHAT ACTUALLY DECIDED THEM

Two things, every time

Whether someone competent looked at what the system produced. And whether that look could stop the thing before it shipped. Neither has a leaderboard. Neither improves when the model gets better. A more capable model produces more plausible output faster, which makes the reviewer's job harder, not easier, because plausible is exactly what slides through a gate nobody is really standing at.

This is why provider diversification, real and worth doing, does not touch the risk that ships broken software. Route across three providers, add the fourth when it launches, and you have changed your cost base and your vendor exposure. Both are genuine. Neither is the control that determines whether what you ship is safe. That control lives at the point where a human either inspects what the system produced or waves it through, and where that inspection can actually halt a release. Change the model as often as you like. If nothing competent is standing at that point, and cannot say no, you have moved the deckchairs and left the hole.

The question was never which model. It is who is checking, and can they say no.



ABOUT THIS NOTE

Where this comes from

This field note accompanies a LinkedIn post on model-provider concentration risk. It sits alongside the Cognitive Delivery Framework, an operating model for building software with AI agents, and the Cognitive Delivery Index, the instrument that measures whether a control demonstrably changes behaviour at the point of action rather than merely being documented. The distance between a control that is written down and one that provably alters what ships is the gap every case above fell through.

SOURCES

Knight Capital: SEC order (2013-222) and Knight Capital trading-loss disclosures, August 2012. CrowdStrike: company Preliminary Post Incident Review and Root Cause Analysis, July to August 2024; Fitch Ratings commentary. Air Canada: Moffatt v. Air Canada, British Columbia Civil Resolution Tribunal, February 2024. Coding-agent database deletion: contemporaneous reporting and the platform's own public postmortem, July 2025. Named cases are used as reported public record; no client or engagement material appears here.

Jason Cruz Falzon · Creator of the Cognitive Delivery Framework and Index

jason@cognitivedelivery.co.uk · cognitivedelivery.co.uk · cognitivedelivery.substack.com