What retail technology leaders are really worried about

Steve Dennis, Darryl Kennedy, Mike Thomas | Spike | 29 June 2026

Two roundtables. Dozens of conversations. One big question about quality in the age of AI. of AI.

What happens to quality when AI is shipping code faster than your teams can review it?

RetailJam is built differently. Festival atmosphere, curated conversations, and retail leaders who turn turn up to talk honestly about what is working and what is starting to break.

Darryl, Mike and Steve hosted two sessions on The Tsunami of Code, and spent the day in the open marketplace talking with retailers at every stage of their journey.

Foundational tech change work is still very much live

AI is not the only thing on people's minds.

Foundational tech change work is still very much live

In the breaks, over lunch, and across the open sessions, the conversations coming to us were grounded and practical. Shopify migrations. PIM overhauls. Integration layer rebuilds. The kind of foundational change work that does not make for a flashy conference headline but defines whether a retailer can move at the pace the business needs.

A significant number of retailers are still working through the platforms, data structures and integration layers that need to be right before anything more ambitious becomes viable.

Transformation came up too. The organisations making real progress are the ones who have stopped treating change as a project with a finish line and started building it as an organisational capability.

WHAT WE HEARD

Foundational change

- AI is not the only thing on people's minds. Plenty of retailers are still working out what platform they are building on.
- Shopify migrations, PIM overhauls and integration layer rebuilds are the work that makes everything else possible.
- Speed to market is a foundation problem before it is an AI problem.
- A significant number of retailers are not AI-first yet. They are still becoming platform-ready.
- New challenges are stacking on top of the old ones, not replacing them.

Shopify migrations are still underestimated

The pressure to move fast on Shopify came up repeatedly, and so did the gap between how straightforward it looks and how difficult it really is.

The front end is the easy part. The back end and the integrations are where speed starts to become risk.

The testing and integration challenges on a Shopify migration are consistently underestimated at the planning stage, and consistently painful in delivery.

> *"The front end is the easy part. The back end and integrations are where speed becomes risk."*

spike

WHAT WE HEARD

Technical debt and foundations

- Debt does not stay still. Every change you build on top of it costs more than the last one.
- The SMEs who understood your legacy systems have left. That is an institutional knowledge problem.
- When your architecture is a black box, you lose the ability to see how bad it actually is.
- "Never waste a good P1" is dark humour. It is also how board investment in remediation gets unlocked.
- The board will fund the exciting stuff. Getting them to fund the foundations is harder.

End-to-end QA, performance and UAT are still the big hurdles

Whatever else has changed, this has not. The parts that break first in production are the same parts that get assumed to work rather than verified: integrations, third-party dependencies, data flows.

UAT in particular remains one of the most painful phases in any programme. The root cause is usually the same: it is treated as a gate rather than a process, and by the time it arrives, there is no room left to act on what it finds.

UAT is treated as a gate rather than a process. By the time it arrives, there is no room left to act on what it finds.

WHAT WE HEARD

Transformation as capability

- Transformation is not a project with a finish line. It is a capability you either build or you do not.
- Ambition and execution capability are not the same thing. The gap between them is where programmes go wrong.
- IT as a blocker is a structural problem, not a personality problem. Fix the structure.
- Performance and transformation are not the same goal. Treating them as if they are is one of the most reliable ways to fail at both.

The roundtables: The Tsunami of Code

How Retail Technology Leaders Protect Quality in an AI-Accelerated World

7 KEY TAKEAWAYS

1 AI is raising the bar and people have questions

The retailers in the room were at every stage of their AI journey, from cautious experiments to agents writing production code.

The questions were strikingly consistent: how do you hold an outsourcer accountable? Who reviews the code the agent wrote? What does go-live even mean when part of your tech was built by a model?

These are governance questions before they are technical ones.

FROM THE ROOM: cost and value

- Nobody has reduced headcount because of AI. Most are growing their teams.
- The value is in speed and confidence, not cost reduction.

2 "It works" is different to "we understand it"

When code is generated faster than anyone can read it, the temptation is to trust the output because it runs. This is exactly the trap.

Code that compiles and passes tests is not the same as code that works reliably in production. The volume has outpaced the review capacity.

FROM THE ROOM: AI and code velocity

- When AI accelerates code production, testing is the first thing that breaks.
- AI is a sledgehammer. Useful. But not precise.
- Code written by an agent today may be hard to support in six months.

3 Testing breaks first

When AI accelerates code volume, the first thing that breaks is testing.

Two patterns emerging:

Pair programmer model
– AI accelerates, human reviews everything.

Agent-connected model
– tickets link directly to repos. Faster. Significantly riskier.

FROM THE ROOM: quality and guardrails

- Humans still need to validate intent, not just syntax.
- Only your lead engineers should be training the model.
- Periodic resets recommended. Agents develop their own patterns.

4 Senior engineers are now the bottleneck

AI does not reduce the need for senior engineering judgement. It increases the demand for it.

Developers are becoming code reviewers and editors rather than writers. You are not a content creator any more. You are a content editor. The volume has tripled. The judgement required has not reduced.

Junior developers are less effective without senior oversight. That is a workforce planning problem most organisations have not yet confronted.

FROM THE ROOM: engineering teams

- AI increases the demand for senior engineering judgement.
- Junior developers are less effective without senior oversight.
- The senior engineer bottleneck is not going away. It is getting tighter.

5 Someone still has to own the code

Speed at the point of generation means very little if accountability gets unclear downstream.

Organisations paying for external delivery should expect measurably more output for the same investment. If a partner is not adopting AI to improve throughput and quality, ask what their long-term roadmap looks like.

"The computer says no" is not an acceptable post-incident response.

FROM THE ROOM: outsourcers

- AI raises the bar for outsourcers. It does not lower it.
- Full outsourcing means losing your internal sense of what good looks like.

6 Data governance is not a solved problem

Conversational analytics tools are only as good as the semantic models beneath them.

In many retail organisations, the same metric is defined differently across finance, ecommerce and marketing. Until that is resolved, self-service AI analytics is not as reliable as it looks in a demo.

Finance uses shipped date. Ecommerce uses order date. They both call it sales. That is not an AI problem. It is a governance problem.

FROM THE ROOM: data and governance

- Conversational analytics are only as good as the semantic models underneath them.
- Self-service AI analytics does not work without investment in unglamorous foundations first.

7

Our Graded Test Approach landed with the right people

For those thinking seriously about AI in their tech stack, we introduced the Spike Graded Test Approach. It sparked exactly the kind of debate we hoped for:

Not "should we test more?" but "how do we test proportionately when the volume and origin of code has changed?"

When a model is part of the build, testing before coding stops being a nice-to-have and starts being the thing that keeps you honest.

READ THE FREE GUIDE

www.wearespike.co.uk/gta

Calibrate to risk

Not everything deserves the same scrutiny. Not everything needs it.

Know where risk sits

What matters is understanding where your risk actually sits.

Test proportionately

Apply the right level of rigour to the right components.

What this all means in practice

Quality is harder to locate

While quality still matters enormously, AI has made it harder to locate. Code can be generated faster than teams can review, test and own it.

Three things protect you

The discipline that protects you is not more speed. It is clear accountability, proportionate testing, and a genuine understanding of where your risk lives.

Not one industry, many bets

This is not one industry uniformly racing toward AI. It is an industry at different speeds, with different foundations. Build the right guardrails for the stage you are actually at.

Get in touch wearespike.co.uk/contact-us