Melissa Perri: AI is a load test that exposes broken product operating models

PM Daily Digest

2026-11-21

Melissa Perri: AI is a load test that exposes broken product operating models

By PM Daily Digest • November 21, 2026

Perri's survey shows AI speeding up delivery but not decisions. Agent buyers are turning operations into GTM, and PMs are prototyping before they write PRDs.

AI speeds up delivery but not decisions

Melissa Perri says the build trap is "a systems problem," not a vision problem. Kodak saw digital coming but wasn't built to act on what it saw [1]. She describes the product operating model as five connected parts: practices and governance, structure and funding, strategy, insights, and culture. Getting one wrong "breaks the entire system" [1]. She calls AI a "load test": building is now so fast that teams can ship far past the point where anyone checked whether the work succeeded [1].

Her spring survey of about 309 product leaders found: - 85% use AI for product work, but only 36% say it makes how they operate stronger [1]. - 62% of PMs say they don't know their company's AI strategy, compared with 19% of the C-suite. Perri reads this as a broken path for getting strategy down to teams [1].

From her time at Athena Health, two concrete fixes: - **Shared labels with sales.** Every roadmap item was tagged experiment, alpha, beta or GA, with agreed rules on what sales could discuss. A prototype that isn't in code can't be sold [1]. - **Different governance.** A company-wide quarterly roadmap review that staff nicknamed "prison" was replaced by monthly delivery meetings inside the product team, plus separate leadership reviews of long-term goals. That moved the conversation from "what will you deliver" to "what did you learn" [1].

Her point is that PMs can't make these changes on their own. They are "system things that affect the entire company" [1].

When the buyer is an agent, operations become GTM

An essay by @SavneetS, which Scott Belsky shared, argues that companies are building the wrong layer when they add a chat widget or copilot. A buyer's agent judges what sits behind the front end: latency, data quality, API surface, how readable your policies are, and how many people stand between a request and a resolution [2]. In its insurance example, a business owner's agent queries nine carriers through structured endpoints and binds a policy in seconds. The carrier that wins is the one whose product "is legible to a machine" [2]. Manual re-entry and approval chains become a "latency tax" that agents price in and avoid [2]. The winners, it says, will have one canonical record per customer and transaction, workflows written as executable logic, and APIs an agent can use without filing a support ticket [2]. Belsky's summary: GTM is increasingly "the sum of your quote latency, uptime, fee transparency, and the machine-readable surface you expose" [3].

Prototype first, write the PRD later

Aakash Gupta relays a point from Madhu Gurumurthy, Head of Product for Gemini apps: vibe coding is becoming a core PM skill. The new sequence is idea → brainstorm with AI → prototype → the team tries it → refine → PRD → build. The PRD still has to pin down non-goals, strategic fit, hypotheses and metrics, and approvals, checks and evals [4]. Gupta also says about half the AI PMs he met at Microsoft Ignite were shipping to production, including some at Microsoft. The other half couldn't [5].

DHH offers a caution on craft. Many vibe-coding projects stall after the satisfying first 80%, leaving rough edges and bugs [6]. He hasn't found a setup where agents get the design right in one shot, so "you need the eye" to steer them [6].

Smaller items

- **Swiggy case study.** A user describes a recurring failure: an order is auto-accepted, a rider arrives, the kitchen is closed, and support cancels 45+ minutes later citing a "technical issue." The user traces it to auto-accept making "accepted" look like "confirmed," with no feedback loop from riders [7]. Proposed fixes: treat a rider's "outlet closed" report as a circuit breaker, require low-reliability outlets to confirm prep has started before a rider is sent, and use mandatory reason codes that trigger actions. The proposed north-star metric is merchant-caused cancellations after rider assignment per 1,000 orders [7].
- **Evals.** Teresa Torres says her hand-written LLM-as-judge prompts keep beating AI-written ones, and asks what she's missing about the enthusiasm for DSPy [8].

- **Push beats pull.** A parent's dashboard for three kids' school platforms got zero traction because it had to be opened. The same connectors, rebuilt into an agent that sends a daily WhatsApp summary, got used [9].
- **Hooks and AI.** Nir Eyal calls the "investment" step of his Hook model a big opportunity right now: use AI or another mechanism to make a product get better the more people use it, through data, content or relationships [10].

Sources

1. Still Escaping the Build Trap: cosa cambia con l'AI? - Melissa Perri
2. X article by @SavneetS
3. X post by @scottbelsky
4. substack
5. How to Ship to Production with AI Agents
6. Why AI Has DHH Working 16-Hour Days
7. r/ProductMgmt post by u/Dry-Pineapple-1809
8. X post by @ttorres
9. r/ProductManagement comment by u/Barnabas2109
10. AMA with Nir Eyal, author of Beyond Belief