

AI Accelerates Building—Product Judgment Now Owns the Rest

PM Daily Digest

2026-08-05

AI Accelerates Building—Product Judgment Now Owns the Rest

By PM Daily Digest • August 5, 2026

The latest PM signals point to a shift from maximizing build speed to managing selection, context, adoption, evaluation, and accountable outcomes.

Big Ideas

AI is exposing the parts of product that code cannot solve. A Product-Tank Auckland talk described the *Makers Manifesto*, created by 45 practitioners and modeled as a starting point on the Agile Manifesto, as an end-to-end view of making: commercial viability, build, adoption, and growth—not build alone. Its principles call for purpose over possibility, explicit context, value measurement, pace matched to customer and GTM readiness, evidence-based learning, and ethical human accountability. As agents absorb build-cycle management, PM leverage shifts toward direction, adoption, and consequence ownership. [1]

Selection is becoming a core product capability. Hiten Shah’s warning is that AI will generate more software, content, designs, and ideas than people can process. His hands-on Ori Eval test—416 calls across more than 20 models—produced different winners for coding, tool use, and vision, leading to a routing policy with guardrails based on how each model failed. The PM question is therefore not just “what can we generate?” but “what deserves to ship, under which quality bar?” [2, 3]

Tactical Playbook

Use an AI shipping gate, not a demo:

1. **Discover the problem.** Ask users what problems they have, not only what features they want, and expect multiple ask-and-build cycles. [4]

2. **Externalize context.** Put customer evidence, domain assumptions, constraints, and success measures where agents consume them; context left in people’s heads will not travel fast enough. EasyVet’s PRDs combine architecture knowledge, clinical expertise, and an actual customer-interview repository. [1]
3. **Evaluate on real work.** Define quality dimensions with domain experts, compare LLM judges with manual evaluations, and report scores as learning signals until validated—not as settled scientific facts. [5]
4. **Delegate reversibly.** Let agents prepare or recommend; require a named owner to approve money movement, deletion, or other hard-to-unwind actions. Log inputs and outputs, a stop condition, an alert path, and recovery steps. [6]
5. **Scale scrutiny to stakes.** When teams shrink, deliberately add perspectives—people or different models—and match review standards to risk. Humans still sign the work. [1]

Case Studies & Lessons

EasyVet used AI to buy back product time—and tie speed to user outcomes. The IDEXX team uses AI and forcing functions to split large “container ship” enhancements into smaller batches, balancing speed with quality and stability to increase learning. Reportedly, reduced PM busywork let a PM spend a week in an Australian clinic, feeding real observations into the PRD. In an ambient medical-notes workflow, the heat map of vets documenting at 8 p.m. moved into standard working hours, with a genuine lunch break appearing. The lesson is to use AI acceleration to increase customer contact and reduce a measurable pain, not merely to produce more artifacts. [1]

Career Corner

For AI PM roles, show lived practice. One AI-PM mock-interview panel explicitly framed its view as non-industrywide but argued that side projects and AI use beyond the day job are becoming more important than scripted STAR-style conflict or negotiation stories. Its indirect questions test empathy, innovation, impact, and whether candidates can identify their own problems; senior candidates are expected to carry work through build, shipping, feedback, and adoption. Build a small AI product and document the problem, metric, evaluation or safety choices, and what user feedback changed. [5]

Tools & Resources

Benchmark model choice on your own work. OpenRouter’s Ori Eval evaluates models against tasks in a codebase; Shah’s test shows why task-specific routing beats a single “best model.” Not Diamond Code is another new router that says it selects model and reasoning effort per step and claims 20–65% lower

costs without quality loss. Treat that savings claim as a hypothesis to test against representative tasks, not a procurement fact. [7, 3, 8]

Sources

1. ProductTank Auckland: A Manifesto for the AI journey
2. X post by @hnshah
3. X post by @hnshah
4. X post by @paulg
5. Principal AI Product Manager (Mock Interview) “How innovative are you?”
6. r/startups comment by u/Visible_Speed8843
7. X post by @OpenRouter
8. X post by @tomas_hk