With AI doing the execution, PM judgment becomes the constraint

PM Daily Digest

2026-11-18

With AI doing the execution, PM judgment becomes the constraint

By PM Daily Digest • November 18, 2026

Rohan Varma describes how OpenAI's Codex team was organized, Aakash Gupta argues taste is now the bottleneck, and practicing PMs report how AI has changed their jobs, for better and worse. Also: designing products for AI agents and validating ideas before building.

How the Codex team ran with two PMs

In Product Compass, Rohan Varma, Cursor's first PM and until recently Product Lead for Codex at OpenAI, describes that team: about two PMs, two designers and 30–40 engineers across roughly 15–20 product areas [1]. They wrote few traditional specs. Sometimes a spec was about ten bullets. Varma set strategic guardrails and engineers made the micro-decisions. Most projects had one or two engineers, because "human-to-human collaboration becomes a big bottleneck" [1]. A brief should set out the problem, outcome, constraints and scope. It should also say what the builder can decide alone and what evidence would change direction. Planning ran four to eight weeks at a time, not six months [1].

The guide also offers two habits for AI PMs: - **Start with the bottleneck.** In his support example, agents struggle to find the right policy, not to write replies. That points to a copilot that drafts cited answers for agents to approve, not a fully autonomous chatbot [1]. - **Diagnose before you re-prompt.** A wrong answer can come from retrieval, from missing context or from generation. Each needs a different fix [1].

The constraint is now judgment

Aakash Gupta opens with a line from Madhu Guru, who led product for Gemini: "I've never felt this behind as a product person." Gupta's argument: engineering capacity used to be the excuse, but a PM can now spin up execution in an afternoon. The hard part is separating signal from noise, deciding what to build and scoping ruthlessly. That team constraint was "a useful governor that prevented bad ideas from shipping" [2]. In a separate note, Gupta says he tested four graph designs against single-shot prompts on 40 PM tasks. The biggest margins came on tasks like PRDs, eval suites and roadmap prioritization. Every winning design worked by checking facts, not by improving judgment, so his advice is to "wire a graph when the failure mode is a wrong fact" [3].

Practitioners split on whether AI has improved the job

An r/ProductManagement thread asked whether PMs still enjoy their jobs. The useful answers fit Gupta's argument: - **The upside.** A PM without analyst support says data analysis went from three days to an hour [4]. Another spotted segment-level conversion patterns in a BI chart and had Claude code a four-arm onboarding test that was live by ten. The segment-specific variant was ahead by that afternoon [5]. - **The cost.** A PM managing several teams says the cognitive load has moved from writing docs to reviewing AI output. Some developers now skip reading the spec, and management wants "more and faster" [6]. - **The exposure.** One commenter argues that documentation was never the job. Now that it's 10x faster, PMs who avoid decisions are "exposed" [7]. Another says experience depends on the setup: empowered PMs may be having their best time, while those in feature factories have it worst [8].

What "AI-native" can mean

Run the Business says most heads of product have been handed an "AI-native" mandate in 2026. It can mean one of three things, or a mix: AI in the product surface, an operating model built around an AI development lifecycle, or a strategy focused on how to win rather than what to ship. Which to pursue "depends" on the company [9].

Designing for agents

Ravi Mehta looks at Descript, which wanted outside AI agents to edit videos without ever opening its interface. That raised a choice: build a separate machine interface, or make one product that serves both humans and agents [10]. He frames designing for agents as designing "for humans — one step removed" [10]. In a related post, Hiten Shah argues AI video needs evals against a quality bar for story, pacing and brand. His team tested AI editors on real footage and plans to share what worked on Friday [11].

Briefly

- **Validation:** In an r/startups thread on vetting ideas, one founder's playbook: ask about the last time the problem happened, not "would you use this?". Do the job by hand for five to ten people. Treat a deposit as worth more than a hundred signups, and competitors as evidence of demand [12].
- **Sharing work:** In a podcast with Petra Wille, Teresa Torres urges PMs to share specific work they are personally doing, even null results such as "we tried 20 things… no efficiency gains" [13, 14].

Sources

1. How to Become an AI PM in 2027: The Complete Guide From Cursor's First PM
2. substack
3. substack
4. r/ProductManagement comment by u/Seranz0
5. r/ProductManagement comment by u/stay_true_to_you
6. r/ProductManagement comment by u/livinglogic
7. r/ProductManagement comment by u/HustlinInTheHall
8. r/ProductManagement comment by u/Bernhard-Welzel
9. What Does AI-Native Mean for a Product Org?
10. When AI agents use your product, who are you designing for?
11. X post by @hnshah
12. r/startups comment by u/matif
13. X post by @ttorres
14. X post by @ttorres