

Disconfirm the hire; specialize the model

Recommended Reading from Tech Founders

2026-08-21

Disconfirm the hire; specialize the model

By Recommended Reading from Tech Founders • August 21, 2026

A concise resource brief built around two high-signal recommendation clusters: an adversarial reference-checking playbook surfaced by Sarah Tavel and Sarah Guo, and Harvey’s domain-specific post-training research update highlighted by Sarah Guo and Aaron Levie.

The strongest new recommendations form a practical pairing: make hiring decisions more adversarial, then tailor AI systems to the workflows where specialization can pay off.

Standout: *How to run a real reference check*

- **Content type / creator:** X thread by @dittycheria.
- **Link:** Read the thread
- **Recommended by:** Sarah Tavel called it “so much wisdom” on references, while Sarah Guo called it “great advice for founders on references.” [1, 2]
- **Key takeaway:** A reference check should try to disconfirm the hire—not validate a decision that has already hardened. The thread recommends asking what you are missing, where the candidate is weaker than they appear, and what you will wish you had known six months later. [3]
- **Framework:** Seek “blind” references from former bosses, peers, direct reports, and customers rather than relying only on candidate-provided names; discount praise and amplify criticism to counter politeness; ask questions that force specific examples; and use forced rankings such as “top 1%,” “top 10%,” or “top 25%.” [3]
- **Why it matters:** This turns references into an adversarial decision tool for senior hiring. Its final test is practical: discover why you may be making the wrong choice while you can still change it. [3]

Technical companion: *Update on our post-training effort*

- **Content type / creator:** Technical X article and research update by Harvey researchers @nikogrupen, @ItsJulioPereyra, @calvincongelado, @vtrengarajan, and @gabepereyra. [4]
- **Link:** Read the article
- **Recommended by:** Sarah Guo described it as a blueprint for the performance and efficiency gains possible from post-training, in-domain data, and workflow understanding. [5] Aaron Levie called it a strong account of how applied-AI companies can lower costs and improve accuracy, while stressing that specialization makes sense only when domain expertise and repeated task volume justify it. [6]
- **Key takeaway:** Harvey reports on Tenet, a Kimi K3 base model post-trained with Fireworks for long-horizon legal work, and characterizes its initial results as promising for performance and cost efficiency. [4]
- **Why it matters:** The useful decision rule is narrower than “post-train everything”: specialize when a team understands an enterprise workflow deeply and has enough similar work for optimization to matter; otherwise, a general-purpose frontier model may be sufficient. [6]
- **Caveat:** This is a first-party Harvey update, so its performance claims are company-reported. Read it primarily for the workflow, evaluation, and training design rather than as independent benchmark validation. [4]

Sources

1. X post by @sarahtavel
2. X post by @saranormous
3. X post by @dittycheria
4. X article by @gabepereyra
5. X post by @saranormous
6. X post by @levie