Should We Engage an AI Deployment Company?

A one-page board brief. Updated 11-08-2026.

The case for

- Embedded engineers arrive staffed and ship to production faster than most internal teams can hire.
- Lab-backed and hyperscaler units bring early access to models and platform credits.
- Purchased or partner-led AI tools succeed more often than internal builds in the MIT NANDA study (about 67% vs 33%; contested).
- Fixed or outcome-based pricing is available from some providers (AWS).

The case against

- Lab-backed partners favor their owner's models, which can mean lock-in.
- Pricing is mostly undisclosed; production unit costs must be negotiated.
- The market is months old: OpenAI's and Anthropic's units launched in 05-2026.
- If the partner owns the know-how, the capability leaves when it does.

Four questions management should answer

1. Which single workflow, with which baseline metric, would we start with?
2. What is the cost per task at our volume, in writing?
3. Who owns the code, evaluations and data at exit?
4. Which model, and what evidence supports it on our task?

Recommended decision

Approve a paid diagnostic of 2-4 weeks with two provider types competing, a written exit plan and a go-live bar agreed before the build.