

Aporen

Seam

The coding agent that asks before it assumes.

Disambiguation-first software engineering for AI-assisted development.

FULLY LOCAL INFERENCE

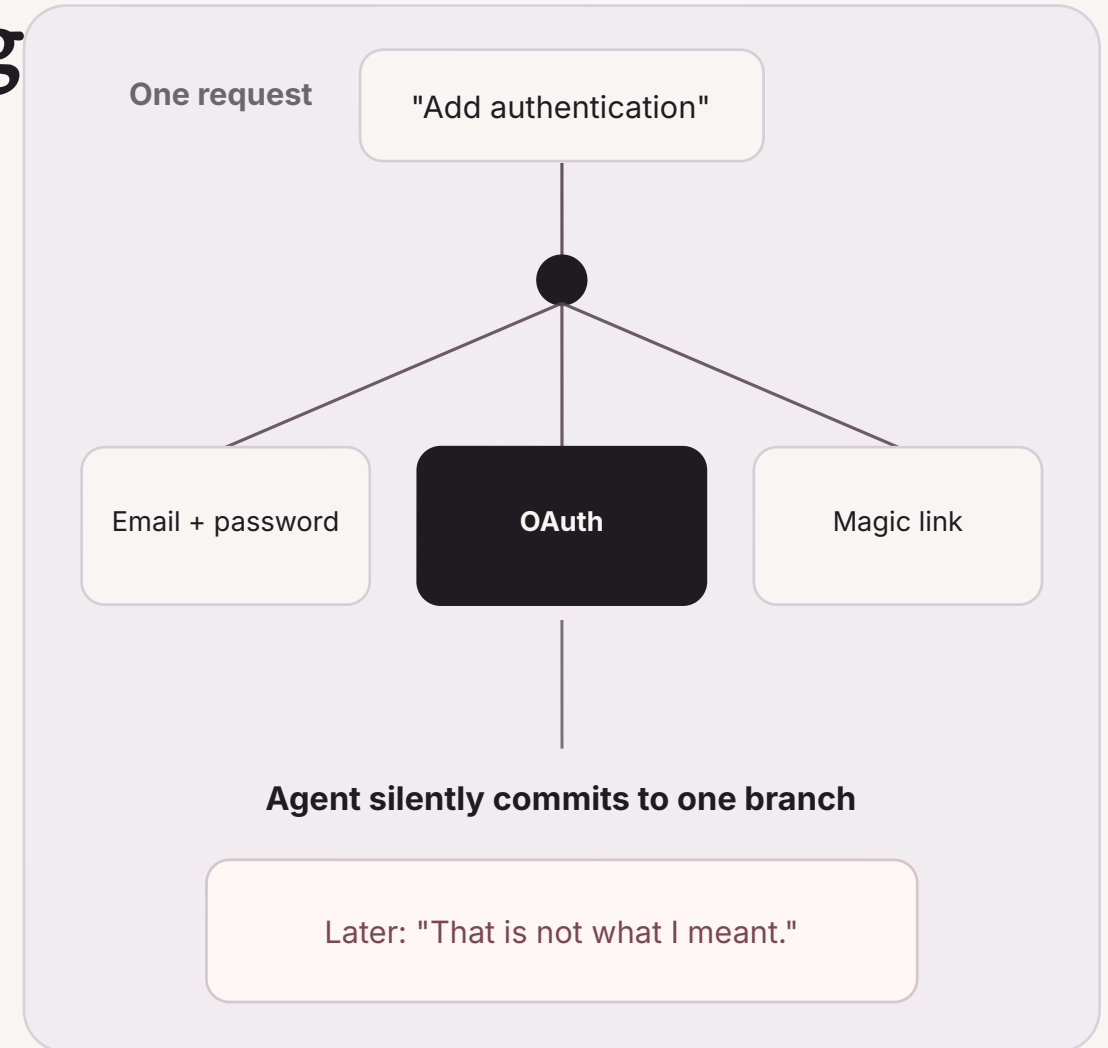
FORKPOINT ENGINE



Coding agents are fast at the wrong interpretation.

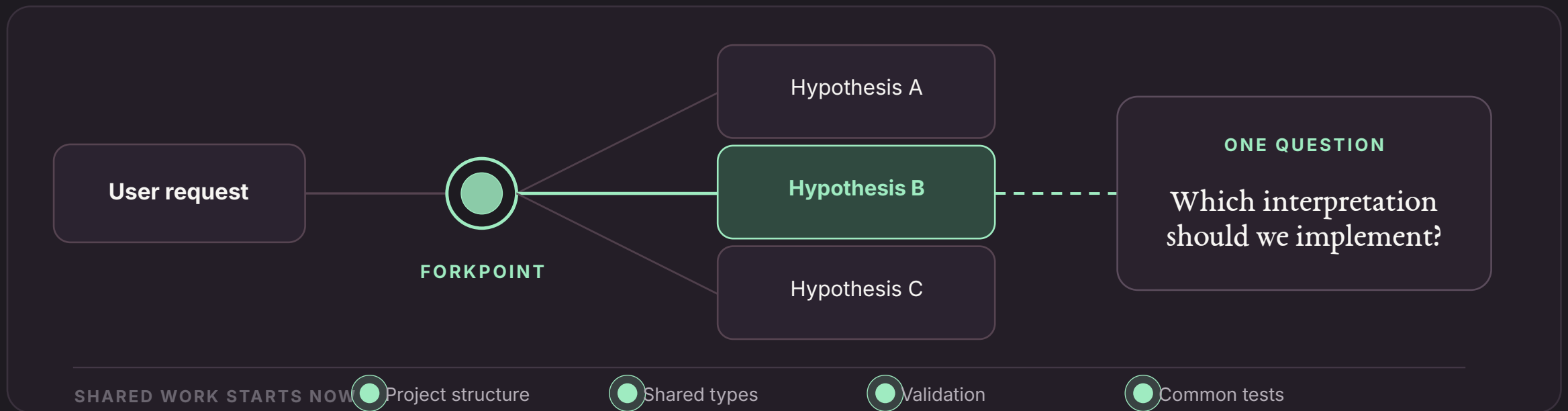
Software requirements are often underspecified. Current agents usually choose one interpretation and proceed with confidence.

- 1 Ambiguity is hidden** The agent turns uncertainty into an unstated assumption.
- 2 Mismatch is discovered late** The developer catches the problem after implementation has already progressed.
- 3 Completed work becomes rework** Time, tokens and trust are lost to the wrong branch.



Seam keeps ambiguity explicit.

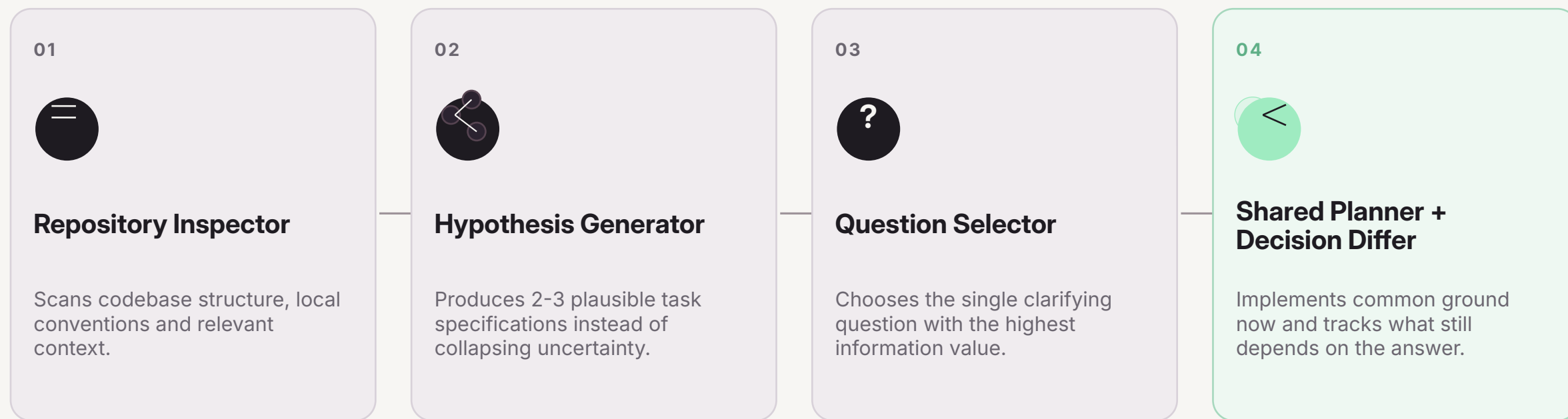
Forkpoint Engine holds multiple plausible interpretations, asks one question that separates them, and starts the shared work immediately.



Clarify only what matters. Build everything else.

Four steps from request to safe execution.

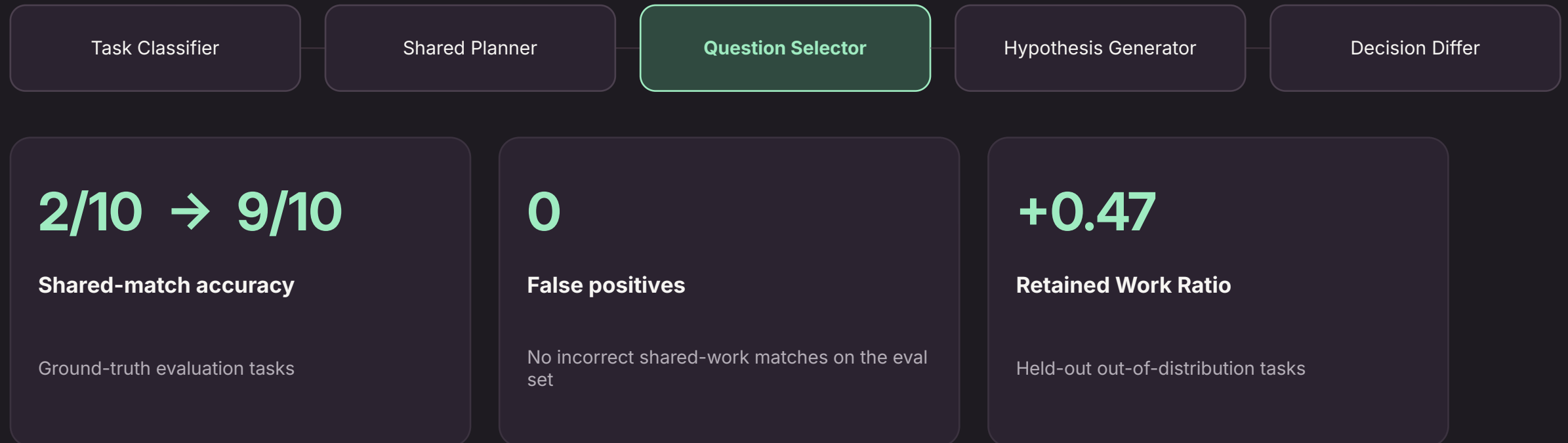
A small reasoning pipeline sits in front of code generation and keeps unresolved choices explicit.



The output is not a guess. It is an implementation with explicit unresolved decisions.

Specialized local inference, measured on held-out tasks.

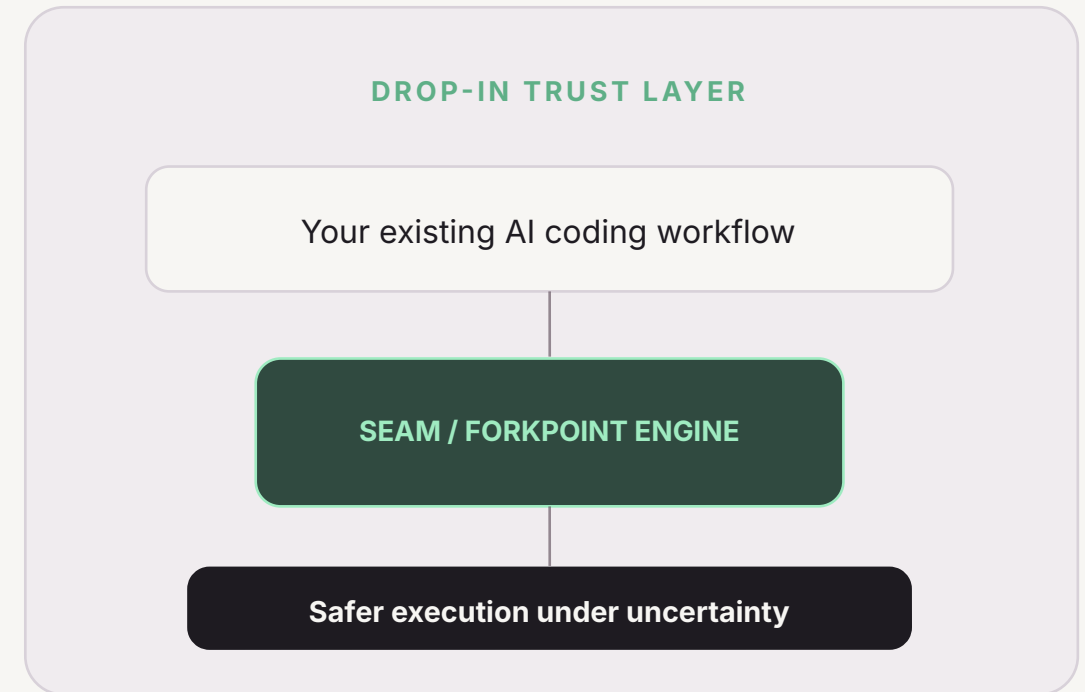
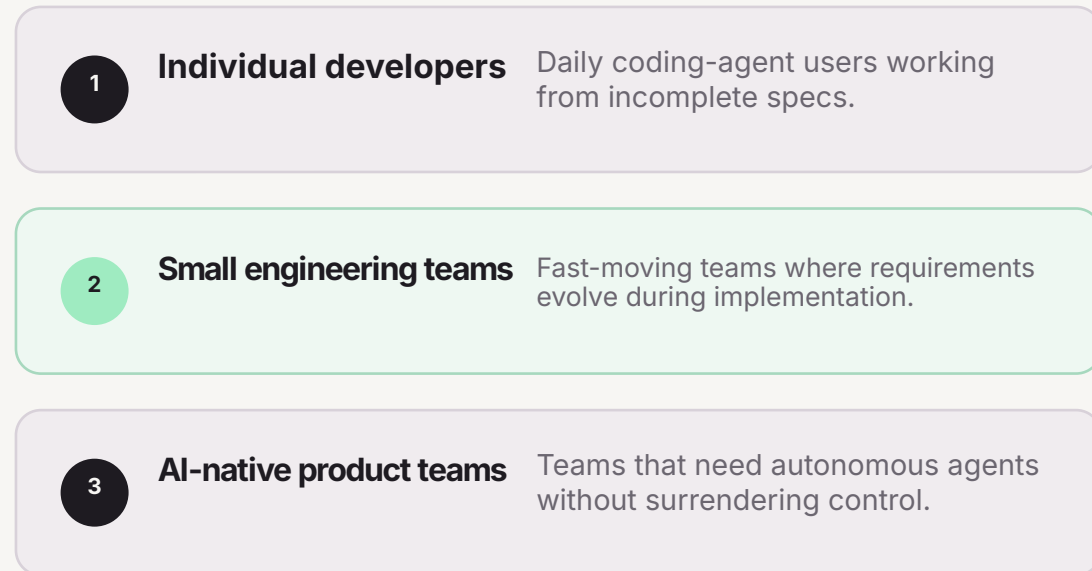
All five core stages run on custom fine-tuned LoRA adapters served through vLLM. No third-party inference dependency in production.



Evaluation discipline Seed-grouped train/eval split • held-out OOD tasks • early stopping

A trust layer for everyday AI-assisted coding.

The initial customer is not "every developer." It is the developer who already uses coding agents and repeatedly pays the cost of wrong assumptions.



Positioning: Seam does not replace the coding workflow. It makes that workflow safer when the specification is incomplete.

One mechanic, expanding domains.

Forkpoint Engine is the reusable layer: keep uncertainty explicit, ask the decision that matters, continue the shared work.

NOW

Seam

Disambiguation-first coding agent

- ☒ Shipping product
- ☒ Local inference stack
- ☒ Evaluation infrastructure

NEXT

Sandbox

Interactive idea validation

- ☐ Real browser research
- ☐ Grounded competitor analysis
- ☐ Forkpoint for product decisions

LATER

Aporen Games

Prompt to playable 3D game

- ☐ Specialized generation models
- ☐ AI playtester with pass/fail checks
- ☐ Autonomous repair loop

Same principle across code, product decisions and game design: preserve useful work while uncertainty is unresolved.

Aporen is building the local reasoning layer for safer coding agents.

ABOUT

Solo-founder AI studio based in Almaty, Kazakhstan.

- Bootstrapped, pre-funding
- Seam in active development
- Self-hosted vLLM + LoRA stack
- Internal benchmarks and evaluation pipeline

WHAT WE ARE ASKING FROM NVIDIA

Help us scale local inference without giving up control.

- 1 **GPU / cloud credits** Scale training, evaluation and local model serving.
- 2 **Technical partnership** Guidance on efficient inference and deployment.
- 3 **Ecosystem access** Introductions to design partners and early technical users.

Seam is shipping. The next constraint is compute, evaluation scale and access to the right technical users.

aporen.ai