

Evals are the scanner. Ship/stop is the layer.

WHY NOW
Everyone tests; next swap needs ship/stop on your workflows.

WHY THIS
Portable ship/no-ship in your CI — not another dashboard score.

89%
obs

52.4%
offline evals

24×11
winner changes

422
OR routes

AppSec	AI ship
Thousands of CVEs / deps	Dozens of models + routers + silent swaps
Can't patch everything	Can't re-eval every workflow on every change
Scanning ≠ prioritized fix	Eval scores ≠ ship/stop on your paths

LangChain N=1,340 · freeze 20260711_215748 · vantageai-private.netlify.app/ship-decision-layer

MODELS · ROUTERS · AGENTS

INSPECT

LangSmith · Braintrust · evals — mainstream

DECIDE — MISSING

Ship/stop on YOUR workflows in YOUR CI

Measurement

Evals

Obs

Ship decision

Ship CI

Owners

89%

89%

∅

∅

W×M×C: hobbyist 2/mo → fleet 4,800/mo · OR catalog accelerating · Qualifier: 3 workflows · CI blocks?