VENDOR ASSESSMENT GRID

The five questions before buying an agent

LEVEL	Operational
DURATION	1h - 2h
USE	Running an agent
TRACK	Toolkit

OBJECTIVE Separate a governable agent from a black box. Score each answer out of two points; the total out of ten informs the go / no-go (← ACF-08).

1 THE FIVE QUESTIONS · GOOD ANSWER VS WARNING SIGN

1 · BELOW THE CONFIDENCE THRESHOLD?

Good: it escalates or stops, it traces, thresholds configurable. Warning: "it carries on" ___ / 2

2 · LOAD AT SCALE?

Good: bottlenecks named, production figures. Warning: silence, "it scales by itself" ___ / 2

3 · A REAL AUDIT LOG?

Good: a signed trace (Ed25519 + SHA-256), queryable. Warning: screenshots ___ / 2

4 · MODEL VERSIONING?

Good: breaking changes announced, testing mandatory. Warning: "it improves by itself" ___ / 2

5 · LATENCY UNDER REAL LOAD?

Good: production figures, named references. Warning: extrapolated demo metrics ___ / 2

2 VERDICT · TOTAL OUT OF 10

8 TO 10

Ready for production

4 TO 7

Push the vendor: demand the missing evidence

0 TO 3

Real governance risk: do not deploy at scale

The sales demo always shines; production is won on these five proofs.

© 2026 Vincent Dorange · acfstandard.com/licence