

Agent Reliability Review: \$12,500, fixed, 4 weeks.

For the VP of Engineering or head of AI whose agent works in demos, while nobody can say how often it fails on real traffic or why.

You receive	A pass rate on a test suite built from your own traces · every failure mode named, counted and reproduced as a runnable test · the suite running in your CI · a written go or no-go with every fix estimated
Timeline	Kickoff, 45 minutes · Day 5: first failure list from 50 traces · Week 3: suite running in your CI · Week 4: decision memo and a 30-minute review
Scope	One agent, one workflow, up to 200 traces or transcripts, 32 delivery hours. Work outside scope by written change order at \$225 an hour.
You provide	Read access to traces, one kickoff, two 30-minute reviews. Modeo writes every test.
Included	A re-run of the suite 30 days after your fixes, with the new pass rate · \$3,750 credited to a Production Build or Monitor and Improve signed within 30 days
Data	Mutual NDA first. Redacted traces, or Modeo works inside your environment on your access.
Payment	50% on signature, 50% on delivery, net 15. Every deliverable is delivered whatever the verdict.
Not included	The fixes themselves, quoted separately · uptime SLAs · healthcare data under HIPAA

See the format first

Modeo ran this method on FamilyCheck, a voice agent it built and operates: 88% scenario pass across 53 scenarios, six failure modes, one safety fix before adding patients. Sample report:

modeo.io/modeo-sample-reliability-review.pdf

Book: cal.com/modeo/30min

Email: hello@modeo.io

Web: modeo.io