

Free API lanes — text and structured, ten attempts each

PUBLIC RESULT · CAMPAIGN SUMMARY

67 passes

Across 124 graded attempts; no cross-exam quality score

COVERAGE

2 exams · 7 lanes

COMPLETED ATTEMPTS

136

MODEL FAILURES

57

LANE ERRORS

12

WITHHELD / INVALID

0 / 0

Results by AI exam

Exam	Model	Passed / attempts	Errors
Record normalization and duplicate QA	@cf/openai/gpt-oss-20b	9/10	0
Intent and entity annotation	@cf/openai/gpt-oss-20b	10/10	0
Record normalization and duplicate QA	openai/gpt-oss-20b	10/10	0
Intent and entity annotation	openai/gpt-oss-20b	1/1 + 9 lane errors	9
Record normalization and duplicate QA	openai/gpt-oss-120b	7/10	0
Intent and entity annotation	openai/gpt-oss-120b	3/3 + 3 lane errors	3
Record normalization and duplicate QA	codestral-latest	9/10	0
Intent and entity annotation	codestral-latest	0/10	0

First 8 of 14 model/exam results in report order. The complete table is linked on page 2. Technical errors are counted in attempts but are not model failures.

4 verified records. Compare models within the same exam and input condition.

What was evaluated

INPUT CONDITION
Conditions differ by exam and model lane; the detailed matrix retains each cell without merging unlike tasks.
EXAM AND REVISION
2 exam revisions; each fingerprint is retained in the detailed report

Limits on the claim

- This summary is accounting, not a leaderboard or a broad capability score.
- Low sample cells support counts only; no ranking is inferred.
- Lane errors, withheld and invalid attempts are retained outside the model-failure denominator.
- Per-exam conditions and quality axes remain in the detailed report.

Evidence

CAMPAIGN	CANONICAL DIGEST
fc-3377c49ce724	4 sealed record digests in the detailed report

Open the full evidence matrix and audit report →

<https://vvdexops.com/reports/fc-3377c49ce724/>

Public disclosure excludes submissions, raw answers, private references, exact geometry, frame timestamps, and hidden-test material.