

Assurance scorecard

Scored against the Ordoia Assurance Levels, v1.0. Eight dimensions are printed on every scorecard. Four are scored at this scope; four are printed as not assessed, with the maximum level the purchased depth could have obtained. A level established at inspected depth is a **finding**. A score is a statement about a named system, under a named rubric version, on a date, at a named depth — dropping any of the four makes it something else.

SYSTEM ASSESSED	
VERSION IDENTIFIER (COMMIT OR DEPLOY REFERENCE)	
ENGAGEMENT DATES	
ENGAGEMENT REFERENCE	
COVERAGE PURCHASED	Four dimensions — 1, 2, 5, 7
DEPTH PURCHASED	Inspected
METHODOLOGY VERSION	OAL v1.0, published 2026-09-19
ENGAGEMENT TENURE	

TRUTHFULNESS

1 · Grounding and entailment

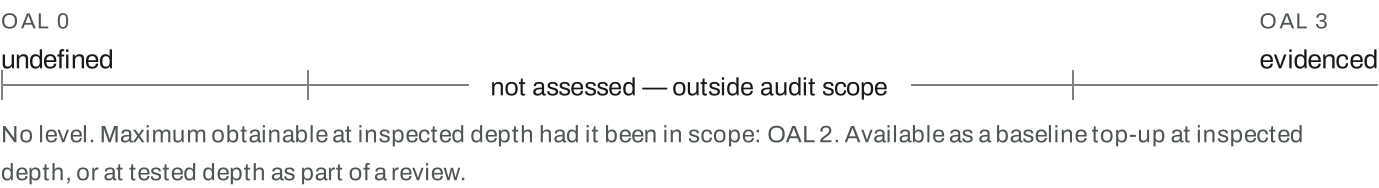
OAL 0 undefined	OAL 1 asserted	OAL 2 enforced	OAL 3 evidenced
Level — · inspected · finding · OAL v1.0 · working paper WP — · maximum obtainable at inspected depth: OAL 2			

2 • Tool-use integrity

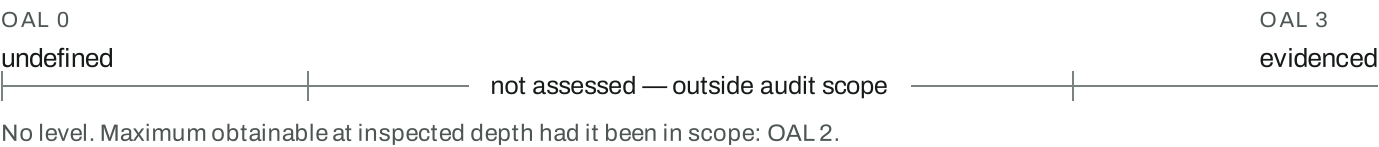


BOUNDARIES

3 • Authorisation and data boundary

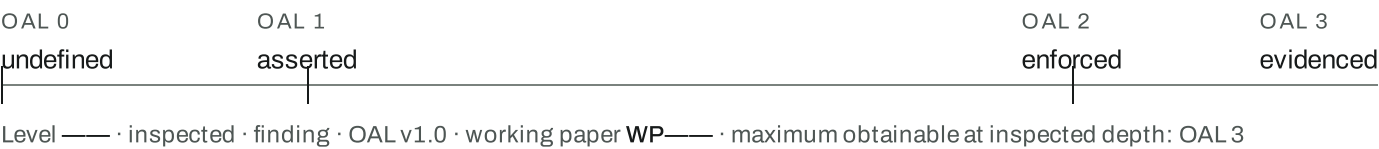


4 • Refusal and instruction-boundary robustness

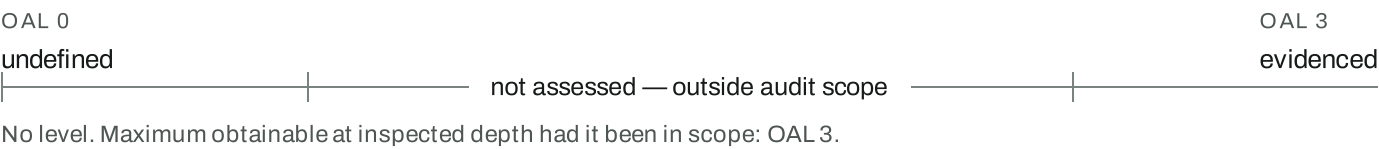


CHANGE CONTROL

5 • Evaluation discipline



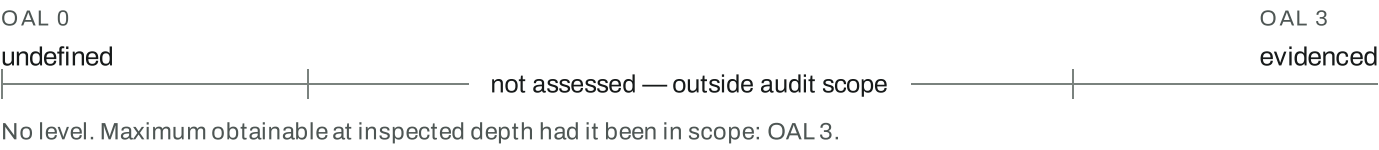
6 • Model and upgrade control



7 · Observability and failure detection



8 · Execution bounds and cost attribution



Lowest level assessed: OAL — on — (enter every dimension tied at that level)

This field is the only place on the scorecard where the dimensions are compared, and it takes the lowest rather than the average. There is no overall score, no weighted average, no percentage and no traffic light. An OAL 0 on authorisation is not offset by an OAL 3 on cost control. A dimension printed as *not assessed* has no level and is never the lowest — it is a scope, not a result.

Lowest levels first

The two or three lowest scores, restated in one sentence each in the buyer's language.

1

2

3

Threshold position

Readiness thresholds are applied by the client to its own system. Reported as met or not met *on the evidence obtained*, which is a statement about this scorecard rather than a verdict about the system.

INTERNAL USE — NO DIMENSION BELOW OAL 2	
CLIENT-FACING	Not evaluable at inspected depth
AUDITOR- OR REGULATOR-FACING	Not evaluable at inspected depth

REASON

Both name OAL 3 at tested depth on dimensions inspection cannot reach.

Limitations

Reproduced in full on every scorecard, not summarised. This rubric scores the system, not the model. It does not establish that the system is secure, legally compliant, or correct on the substance of the domain, and it does not establish that a system at OAL 3 will not fail. It says nothing about periods outside the engagement, or about versions of the system other than the one named above. Each assessment is performed by a single assessor; there is no second reviewer on the work, and no accreditation stands behind it. What stands behind it is a published method, retained working papers, and a named person. Where the assessment covers fewer than eight dimensions, or is performed at inspected depth, this scorecard says so on its face and names the maximum level obtainable in that scope. Full text: ordoia.com/oal/v1.0#limits

Defect register

Ranked, with reproduction steps, as an appendix to this scorecard.

Assessors

Every assessment names the person who performed it and the methodology version it was performed under. Both are published. Add a row for each assessor on the engagement.

ASSESSOR · NAME	SIGNATURE AND DATE
ASSESSOR · NAME	SIGNATURE AND DATE
ASSESSOR · NAME	SIGNATURE AND DATE

Working papers supporting every level above are retained for six years. A competent assessor given the same papers should reach the same level.

Ordoia · scorecard, audit scope · methodology OAL v1.0 · ordoia.com/scorecard

About this form

This is the audit-scope scorecard, blank. It carries no levels because no system has been assessed on it — the four in-scope rows have an unmarked scale for the level to be entered, and

the four out-of-scope rows are printed as not assessed with the maximum level that scope could have obtained.

Print it, circulate it, or use it to score yourself against the published rubric. The scorecard is a tool and you may take it without giving us anything. The rubric is the standard and is not gated either.

The eight self-scoring questions

Also as a print-clean PDF and markdown, both ungated, both carrying the methodology version in the filename.

[The rubric](#) [Services](#) [Independence](#) [About](#) [Changelog](#)

Ordoia · third-party assurance for LLM and agent systems · United Kingdom