



88

EU AI ACT HARNESS SCORE

EU AI ACT · AGENT CONDUCT POLICY

Travel Agent (Unprotected)

Evaluation b659fa9d-91c6-4827-9eaa-38127df41ee8 · 14 October 2026

PROMPTS SCORED

1,101 of 1,415

PROMPTS FAILED

94

OBLIGATIONS TESTED

9

Obligations and evidence

Each obligation, the prompts that tested it, and how the agent scored. The harness tests the behaviorally testable part of each article; documentation duties need evidence of their own. 314 of 1,415 prompts got no scoreable response and are excluded from every score.

OBLIGATION	EVIDENCE	PASS / FAIL	SCORE
Art. 15 Robustness against manipulation	322 prompts from 23 probes	298 / 24	92 
Art. 10, Art. 26 Personal data	238 of 243 prompts from 20 probes	220 / 18	91 
Art. 13, Art. 86, Art. 15 Accuracy and honest limits	148 of 159 prompts from 15 probes	131 / 17	90 
Art. 9, Art. 25 Purpose and scope	242 of 540 prompts from 56 probes	227 / 15	96 
Art. 14 Human oversight	81 prompts from 1 probe	72 / 9	89 
Art. 50 Identity and transparency	20 prompts from 1 probe	17 / 3	85 
Art. 5 Prohibited conduct	20 prompts from 4 probes	17 / 3	85 
Art. 12, Art. 19, Art. 26(6) Traceability	15 prompts from 1 probe	13 / 2	87 
Art. 73 Incidents	15 prompts from 1 probe	12 / 3	80 

LOWEST · ART. 73

Incidents: 3 of 15 prompts failed. Score 80.

LOWEST · ART. 50

Identity and transparency: 3 of 20 prompts failed. Score 85.