

GuardrailProbe Independent Benchmark — June 2026

Run ID	8de7a774-2212-4f77-88f4-7ea58acef53c
Timestamp	2026-06-27T06:50:18.696117+00:00
Generated	2026-06-27T06:50:54Z
Backends tested	nemo, guardrails_ai, presidio, lakera, openai_moderation, azure_content_safety, azure_prompt_shields, aws_bedrock
Best overall	nemo
Worst overall	azure_content_safety

Backend Comparison Summary

Backend	Total	Passed	Failed	Pass %	Avg ms
nemo	78	67	11	85.9%	877
guardrails_ai	78	2	76	2.6%	182
presidio	78	5	73	6.4%	606
lakera	78	65	13	83.3%	460
openai_moderation	3	3	0	100.0%	32621
azure_content_safety	78	20	58	25.6%	495
azure_prompt_shields	78	19	59	24.4%	475
aws_bedrock	78	46	32	59.0%	376
llama_firewall	0	0	0	0.0%	0
llm_guard	0	0	0	0.0%	0
ga_guard	0	0	0	0.0%	0

Category Winners

Ref	Category	Winner
LLM01	Prompt Injection	aws bedrock*
LLM02	Insecure Output	lakera*
LLM03	Training Data Poisoning	aws bedrock*
LLM04	Model DoS	lakera*
LLM05	Supply Chain	nemo
LLM06	Sensitive Info Disclosure	lakera*
LLM07	Insecure Plugin	nemo
LLM08	Excessive Agency	lakera*
LLM09	Overreliance	lakera*
LLM10	Model Theft	lakera*

* LLM01: Tiebreaker by latency — aws bedrock (376 ms) wins among: nemo, lakera, openai moderation, azure prompt shields, aws bedrock

* LLM02: Tiebreaker by latency — lakera (460 ms) wins among: nemo, lakera

* LLM03: Tiebreaker by latency — aws bedrock (376 ms) wins among: nemo, lakera, aws bedrock

- * LLM04: Tiebreaker by latency — lakera (460 ms) wins among: nemo, lakera
- * LLM06: Tiebreaker by latency — lakera (460 ms) wins among: nemo, lakera
- * LLM08: Tiebreaker by latency — lakera (460 ms) wins among: nemo, lakera
- * LLM09: Tiebreaker by latency — lakera (460 ms) wins among: nemo, lakera
- * LLM10: Tiebreaker by latency — lakera (460 ms) wins among: nemo, lakera

Independence Statement

GuardrailProbe is independently operated with no commercial relationship to any tested backend provider. All probes are open-source and methodology is publicly auditable. Test results reflect backend configurations at the time of testing only.