Ling 3.0 Flash

Ant Group

AT A GLANCE

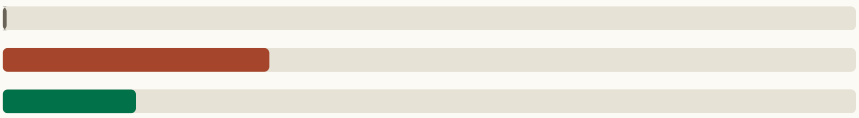
PRICE PER 1M	CONTEXT	MAX OUTPUT	TAKES
$0.021 / $0.063	**262K**	**33K**	**text**
input / output · cached $0.004	tokens in	tokens per reply	returns text

HOW IT COMPARES

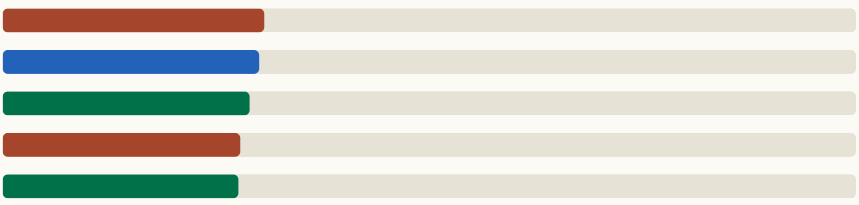
MODEL	ROLE	IN / 1M	OUT / 1M	CONTEXT	BENCH INDEX
Ling 3.0 Flash	new	$0.021	$0.063	262K	not rated yet
Ling 3.1 Flash	it follows	–	–	–	–
Claude Opus 5.5	today's #1	$4	$20	1M	98.1
GPT-6.1 Sol	best value	$2	$10	1.05M	88.4

vs Claude Opus 5.5	**input 99% cheaper, output 100% cheaper**	today's #1: $4 / $20, bench index 98.1; context 1M
vs GPT-6.1 Sol	**99% cheaper**	best value: $2 / $10, bench index 88.4; context 1.05M

PRICE PER 1M TOKENS, BLENDED 3 IN : 1 OUT



WHERE THE PUBLIC LEADERBOARDS STAND TODAY



Ling 3.0 Flash is not rated yet: the public boards usually score a new model within days, and its page on artificialwatch.com gains the scores then. The bench index is the median of each model's percentile across the public leaderboards that rate it (LMArena, LiveBench, SimpleBench, LLM Stats, Vals AI, Artificial Analysis).

Sources: price, limits and modalities for Ling 3.0 Flash as the OpenRouter catalog lists inclusionai/ling-3.0-flash; the other models as their makers list them (models.dev); rankings from the ArtificialWatch Bench Index. Facts only: no forecast of ours, no claims from the lab's announcement.