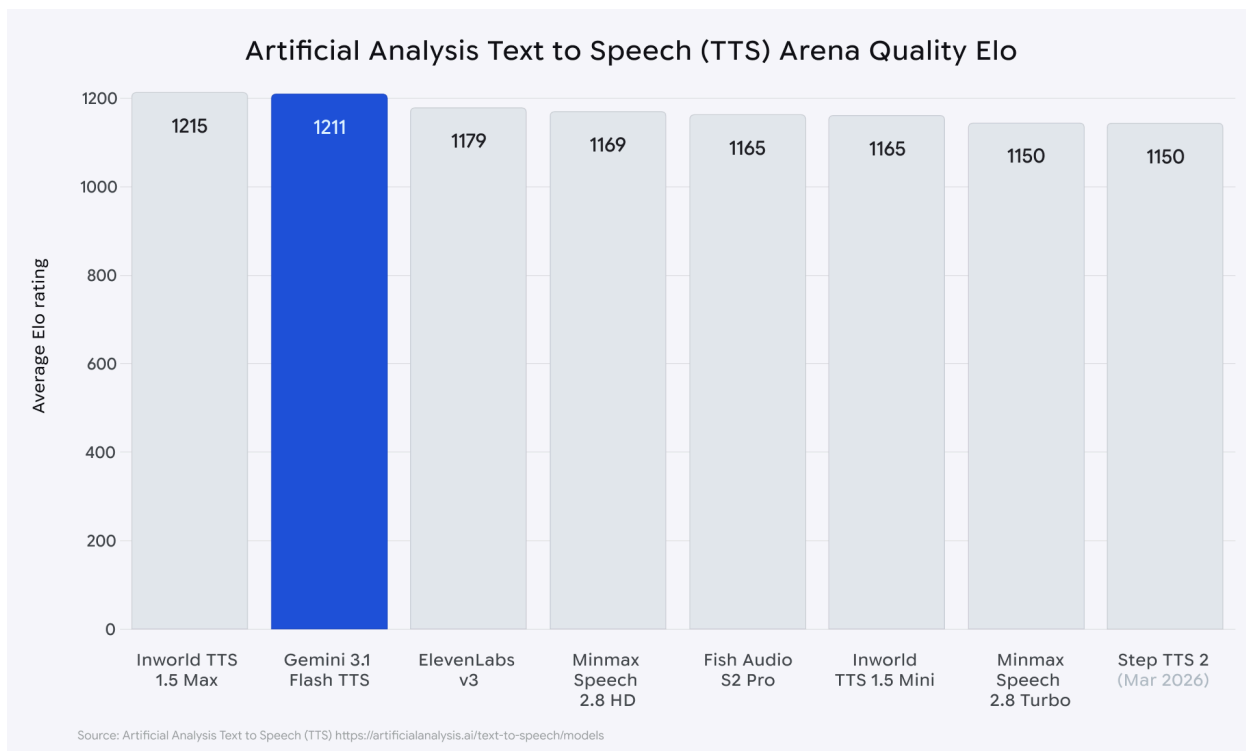


Model Evaluation – Approach, Methodology & Results

Gemini 3.1 Flash Text-to-Speech (TTS)

Approach: Gemini 3.1 Flash TTS was evaluated across a range of benchmarks, including [Artificial Analysis Text to Speech Evals](#).

Results: Gemini 3.1 Flash TTS is now our most natural and expressive text-to-speech model yet, featuring significantly enhanced speech quality. It recently earned a **Elo score of 1,211** on the [Artificial Analysis TTS leaderboard](#), a benchmark based on thousands of blind human evaluations.



Gemini 3.1 Flash TTS delivers top-tier quality (ELO ~1210) at ~\$30/1M chars — roughly 5x cheaper than ElevenLabs v3 at comparable quality. While Inworld leads on raw ELO, Gemini is the only model in the "most attractive quadrant" that also offers 200+ expression tags, natural language control, and native multi-speaker as native model capabilities.

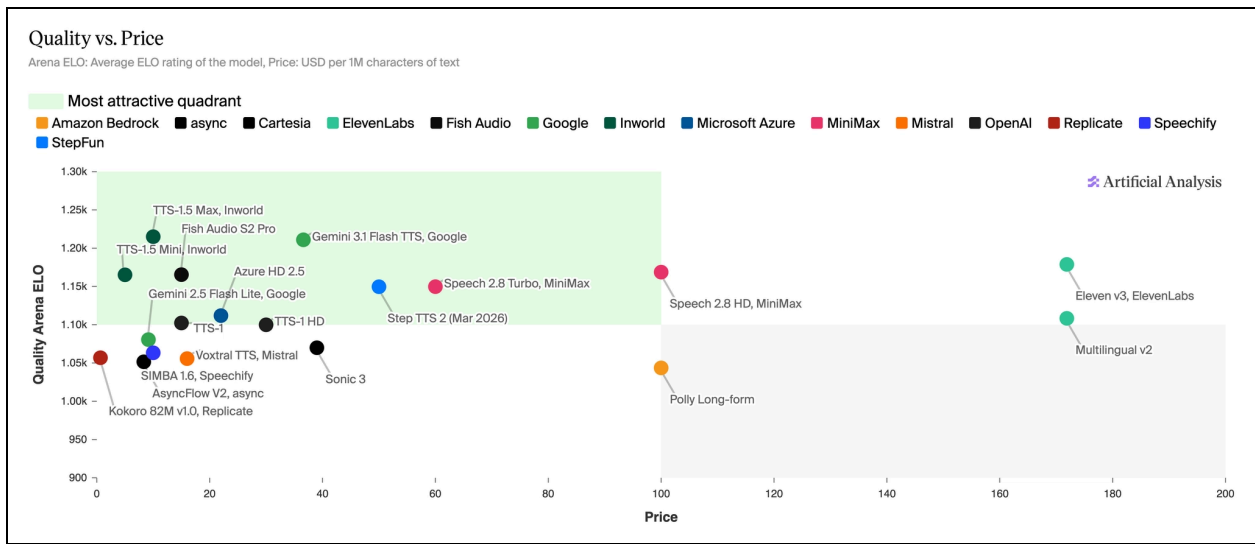


Image source: Artificial Analysis