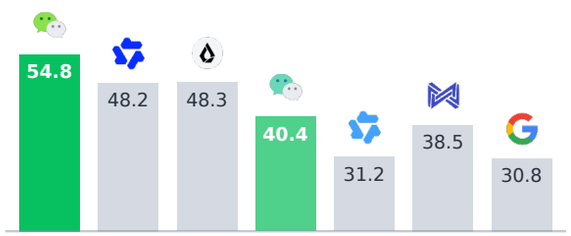
General Benchmarks · WeMM-Ecom-2B & 0.8B vs. Lightweight Peers

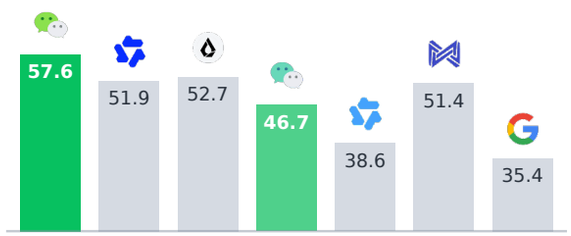
Thinking Mode

Overall (64) · WeMM-Ecom-2B: 63.50 · WeMM-Ecom-0.8B: 52.31

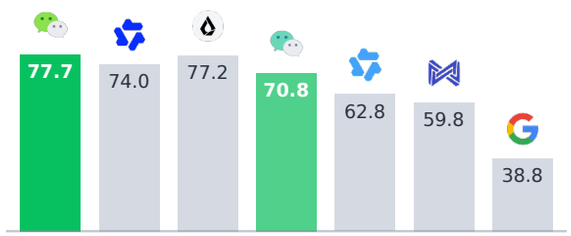
STEM Reasoning
12 Benchmarks



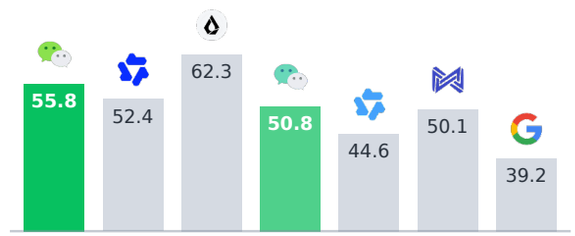
General VQA
11 Benchmarks



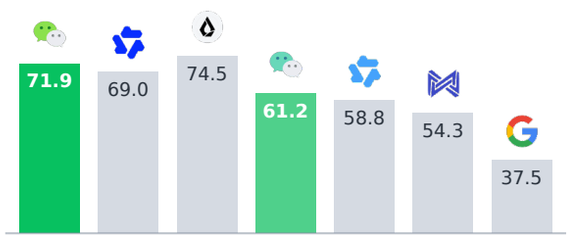
Document & Chart
11 Benchmarks



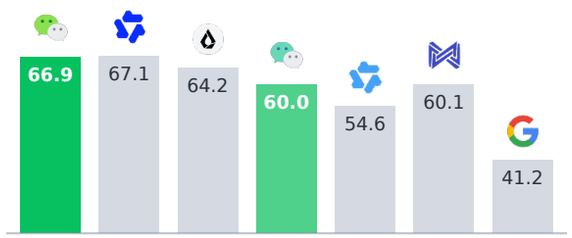
2D / 3D Spatial
6 Benchmarks



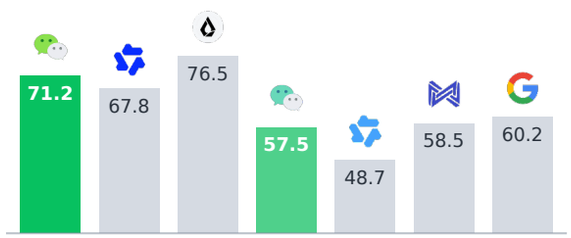
Grounding & Counting
8 Benchmarks



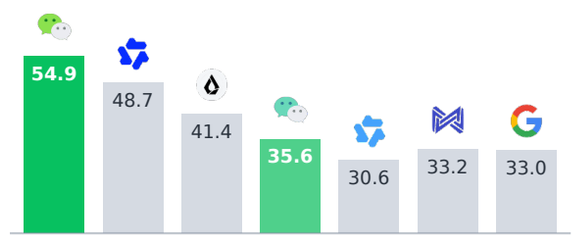
Multi-Image
4 Benchmarks



Instruction Following
5 Benchmarks



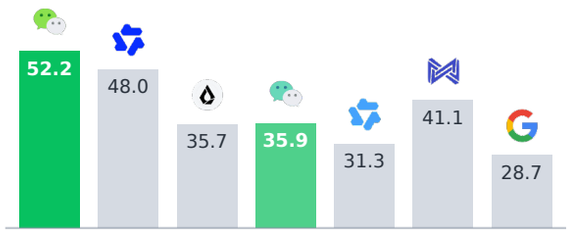
Text-Centric
7 Benchmarks



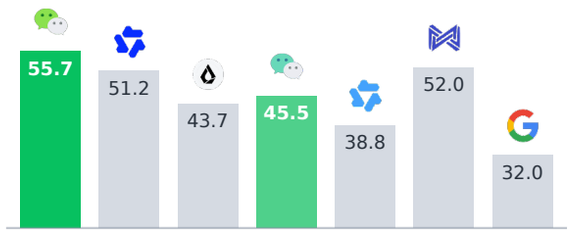
Instruct Mode (Non-Thinking)

Overall (64) · WeMM-Ecom-2B: 60.44 · WeMM-Ecom-0.8B: 48.04

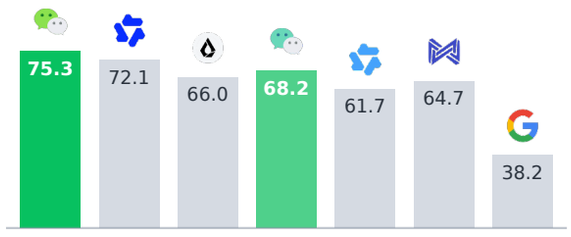
STEM Reasoning
12 Benchmarks



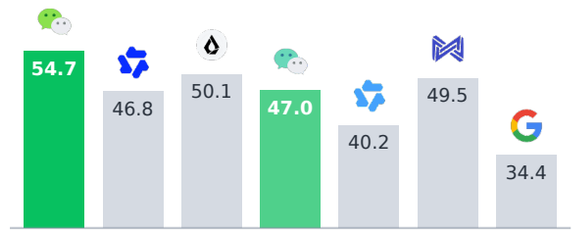
General VQA
11 Benchmarks



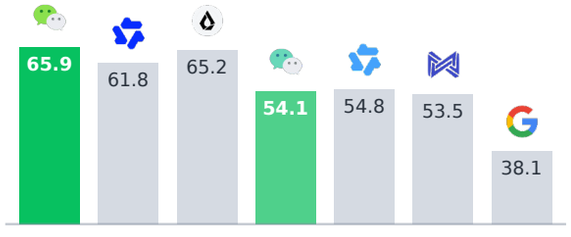
Document & Chart
11 Benchmarks



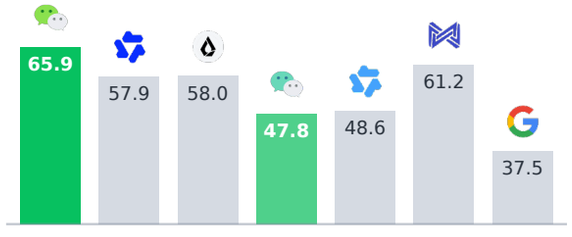
2D / 3D Spatial
6 Benchmarks



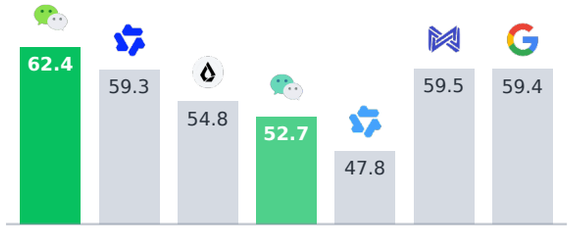
Grounding & Counting
8 Benchmarks



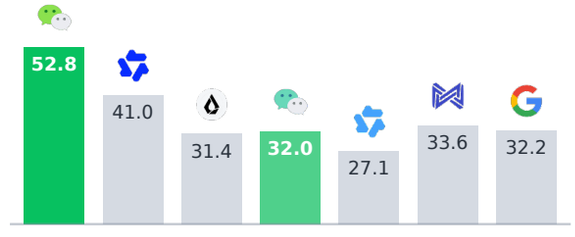
Multi-Image
4 Benchmarks



Instruction Following
5 Benchmarks



Text-Centric
7 Benchmarks



WeMM-Ecom-2B · Qwen3.5-2B · LFM3-VL-2B · **WeMM-Ecom-0.8B** · Qwen3.5-0.8B · MiniCPM-V-4.7-1B · Gemma4-E2B-it

All scores are group averages across the 64-benchmark general suite.