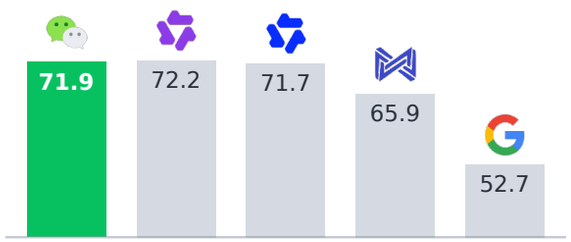
General Benchmarks · WeMM-Ecom-35B-A3B vs. Flagship Peers

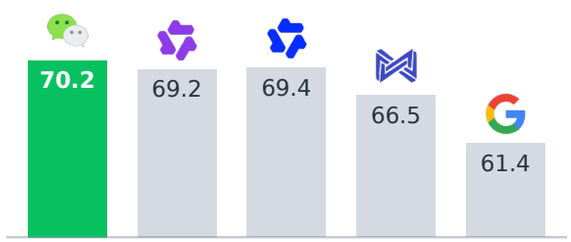
Thinking Mode

Overall (64) · WeMM-Ecom-35B-A3B: 76.37

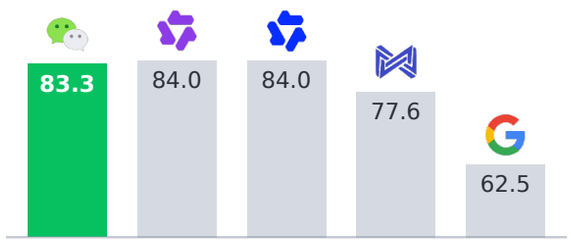
STEM Reasoning
12 Benchmarks



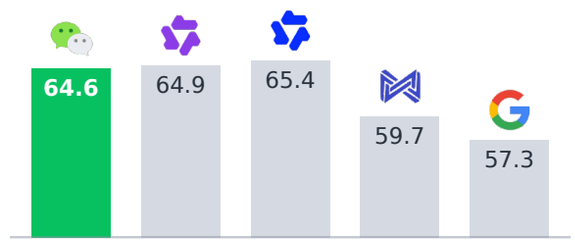
General VQA
11 Benchmarks



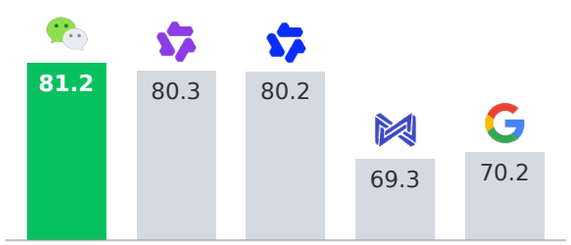
Document & Chart
11 Benchmarks



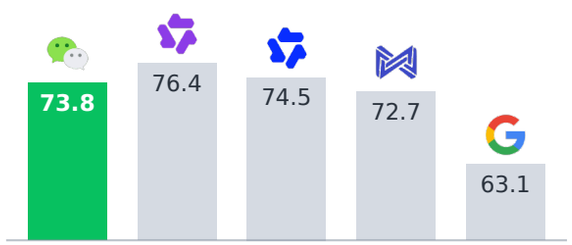
2D / 3D Spatial
6 Benchmarks



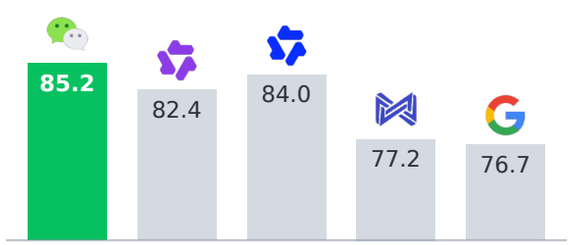
Grounding & Counting
8 Benchmarks



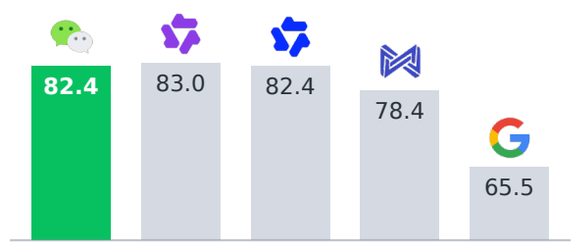
Multi-Image
4 Benchmarks



Instruction Following
5 Benchmarks



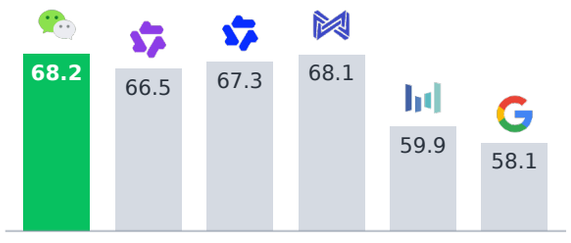
Text-Centric
7 Benchmarks



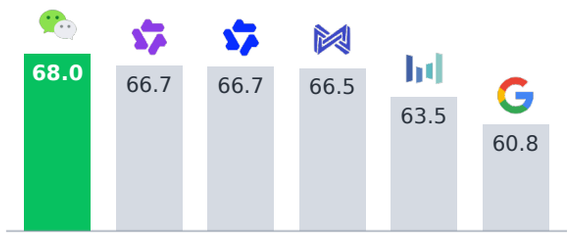
Instruct Mode (Non-Thinking)

Overall (64) · WeMM-Ecom-35B-A3B: 72.58

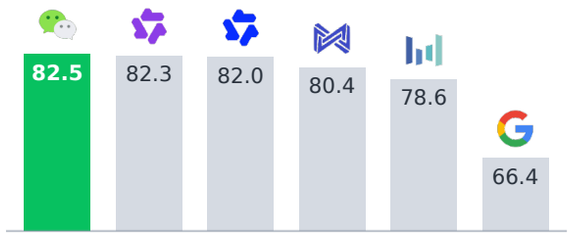
STEM Reasoning
12 Benchmarks



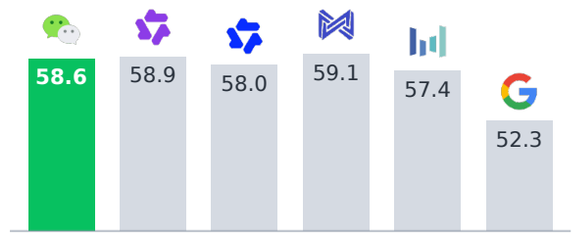
General VQA
11 Benchmarks



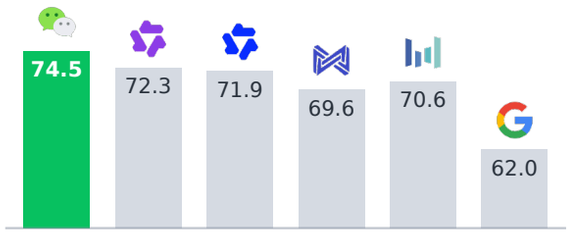
Document & Chart
11 Benchmarks



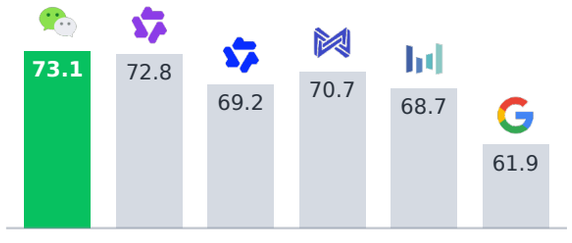
2D / 3D Spatial
6 Benchmarks



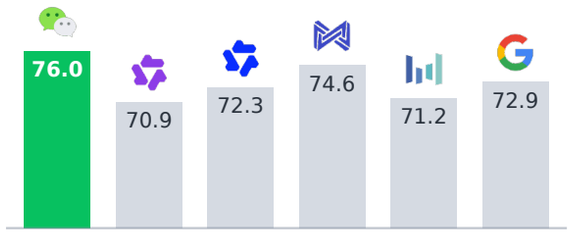
Grounding & Counting
8 Benchmarks



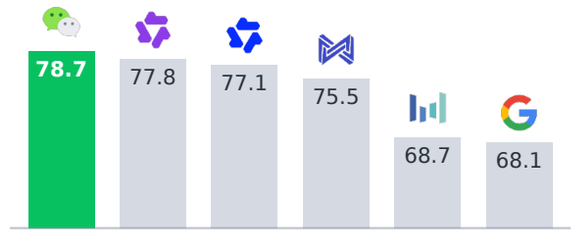
Multi-Image
4 Benchmarks



Instruction Following
5 Benchmarks



Text-Centric
7 Benchmarks



WeMM-Ecom-35B-A3B · Qwen3.6-35B-A3B · Qwen3.5-35B-A3B · MiniCPM-V-4.7-35B-A3B · Valley3-32B* · Gemma4-26B-A4B-it

* Valley3-32B is evaluated in Instruct mode only. All scores are group averages across the 64-benchmark general suite.