

# MCP ATLAS

Evaluating real-world tool use through the Model Context Protocol (MCP)

#	MODEL	PROVIDER	SCORE
1	Muse Spark	meta	82.2
1	claude-opus-4-7 (max)	anthropic	79.1
1	gemini-3.1-pro-preview (high)	google	78.2
2	claude-opus-4-6 (max)	anthropic	76.8
2	glm-5p1	zai	75.6
2	gpt-5.5 (xhigh)	openai	75.3
5	gpt-5.4 (xhigh)	openai	70.6
5	gemini-3-pro-preview	google	70.3
6	claude-opus-4-5 (high)	anthropic	69.8
7	claude-sonnet-4-6	anthropic	69.5
7	gpt-5.2 (xhigh)	openai	67.6
9	kimi-k2p5	kimi	64.4
11	gemini-3-flash-preview	google	62
12	claude-sonnet-4-5 (thinking)	anthropic	59.5
13	glm-4p7	zai	58.1
13	gemini-3.1-flash-lite (high)	google	57.1
13	gpt-5.4-mini (xhigh)	openai	56.7
18	gpt-5.1 (high)	openai	50.1
18	o3-pro	openai	44.5
19	claude-haiku-4-5	anthropic	40.2