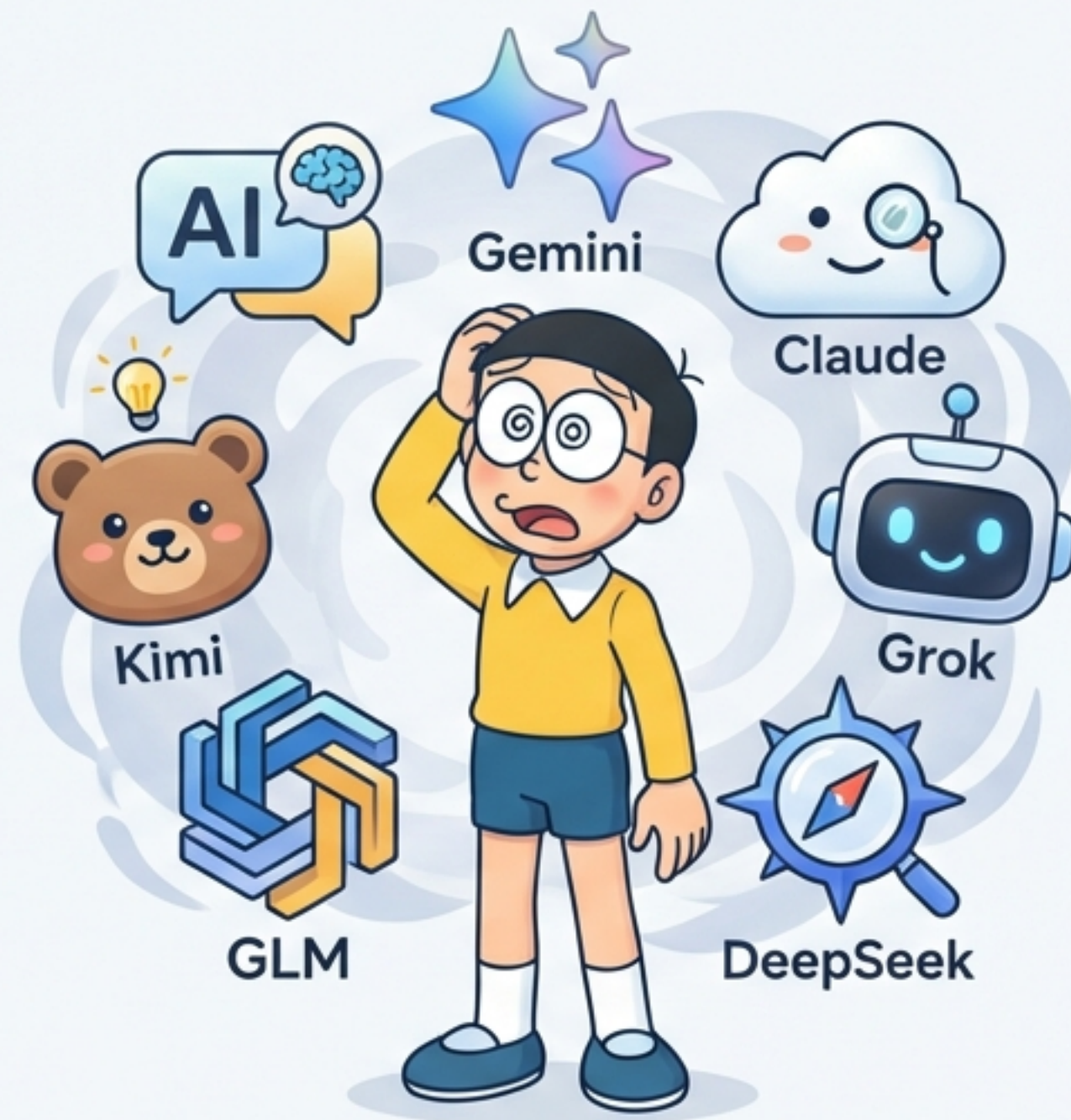


哆啦A梦的AI百宝袋：未来模型关系大揭秘

一张图看懂ChatGPT，Gemini，Claude和它们的“家族”



“大雄，别再烦恼啦！AI的世界看起来就像一团乱麻，但只要有了我的**神奇道具**，一切都会变得清晰起来。我们一起出发，去探索这些**聪明的大脑**背后到底是谁吧！”



AI模型大混战，谁是谁的朋友？谁又是谁的手对手？

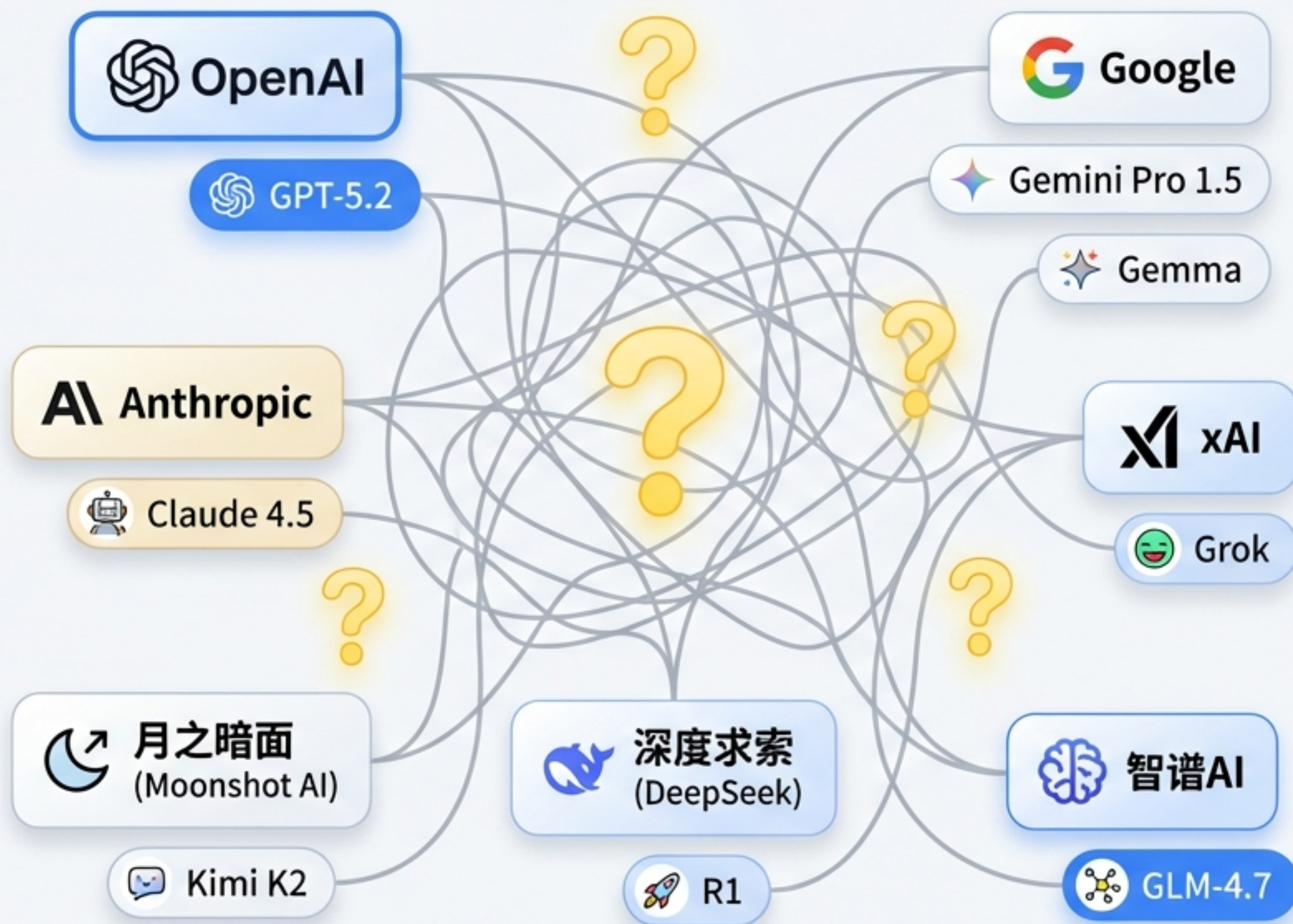


“ChatGPT-5.2 和 5.1 有什么区别？”

“Kimi 和 Claude 哪个更适合写代码？”

“DeepSeek 为什么会引发‘斯普特尼克时刻’？”

“这么多模型，它们的公司之间又是什么关系？”



“别急，大雄！这些问题正是我们要解决的。我带来了一个来自 22 世纪的道具——‘未来AI关系图谱’！它会帮我们把这些一切都梳理清楚。”



神奇道具登场：我们的“未来AI关系图谱”！



让我们一起把它填满，从硅谷的巨头们开始。



“这张图谱就是我们的探险地图！我们会先在‘硅谷巨头’的大陆上标记出那些最著名的然后再去探访‘东方新星’的神秘力量。准备好了吗？”



第一站：探访硅谷的AI巨头家族

OpenAI：生成式智能的开拓者。

Anthropic：专注于构建可信赖、可引导AI的原则派。



Google：拥有强大基础模型架构的科技巨擘。

xAI：致力于理解宇宙、追求真相的挑战者。

“看，这就是硅谷最强大的四个AI家族！他们各自占据着一片领地，并且都在努力创造出最聪明AI。我们先去拜访OpenAI家吧！”



OpenAI的王国：从GPT-4到GPT-5.2的进化之路



关键能力



DALL·E 3: 强大的图像生成能力。



Codex: 为开发者而生的代码智能。



Pulse: 主动为你进行异步研究。



重要的“盟友”

微软（Microsoft）是OpenAI最重要的战略合作伙伴，提供了大量的资金和算力支持。

聚焦旗舰：GPT-5.2的三种强大形态



“最新的GPT-5.2就像一个工具箱，里面有三件针对不同任务的法宝！”

GPT-5.2 对比

形态 (Model)	GPT-5.2 Instant	GPT-5.2 Thinking	GPT-5.2 Pro
定位	日常工作和学习的“快马”	解决更复杂工作任务的“思考者”	应对高难度问题的“最强大脑”
擅长领域	信息查询、技术写作、翻译	电子表格格式化、金融建模、编程	复杂领域编程，错误率更低

知识截止日期

2025年8月



SWE-bench Verified (软件工程基准)

GPT-5.2 Thinking 获得了 **80.0%** 的高分，展示了其强大的代码能力。



Pro用户现在可以选择不同的‘思考时间’ (Thinking Time)，在速度和深度之间做出权衡。（来源：ChatGPT Release Notes, Sep 17, 2025）

两大巨头的不同路径：Google的全能宇宙与Anthropic的原则AI

Google的全能宇宙

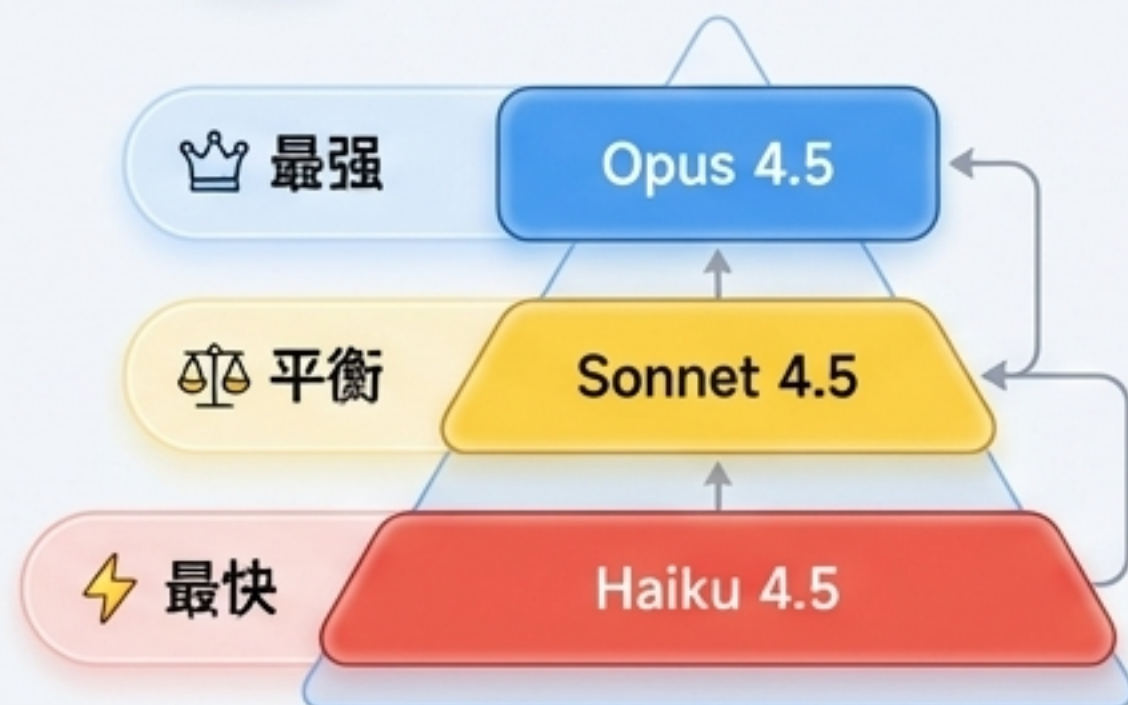


最新模型：Gemini 3（2025年11月发布）

核心能力：强大的原生多模态、高级推理和智能体（Agent）能力。

愿景：成为“XR世界的安卓”，一个灵活、强大、由AI驱动的平台。

Anthropic的原则AI



核心理念

基石：“宪法AI（Constitutional AI）” - 将安全和合规性融入模型的核心决策过程。

重要盟友



Microsoft



NVIDIA

解读Claude 4.5：为不同任务量身打造的三驾马车



“Anthropic的Claude 4.5家族同样分工明确，每个成员都有自己的‘超能力’！”

Claude 4.5 系列对比

模型 (Model)	Claude Opus 4.5	Claude Sonnet 4.5	Claude Haiku 4.5
核心优势	 最智能 ：拥有最高的综合能力	 编程与智能体最佳 ：专为复杂代码和自动化任务设计	 速度最快 ：兼具近乎前沿的性能和极致的速度
独特功能	支持“努力度 (Effort)”参数，平衡思考深度与成本	超强的自主操作能力，可独立工作数小时	首个支持“扩展思考 (Extended Thinking)”的Haiku模型
上下文窗口	200k Tokens	200k / 1M Tokens (beta)	200k Tokens



上下文感知 (Context Awareness)

模型能追踪剩余的上下文窗口，更有效地执行长任务。



编程工具调用 (Programmatic Tool Calling)

显著降低多工具协作的延迟和Token消耗。

硅谷的“异数”：xAI与它连接现实世界的模型Grok



🏢 • 母公司：xAI

🤖 • 旗舰模型：Grok

⚡ • **独门秘籍**：能够利用来自X平台的“实时数据流”，使其具备了对当下事件的感知能力。

🛒 • **重大事件**：2025年3月，xAI通过全股票交易收购了X平台，获得了约6亿月活用户的庞大数据源。



“Grok就像是拥有了一个能随时听到全世界最新消息的顺风耳，这让它在分析市场趋势和热点话题时非常特别！”

第二站：东方新星的崛起



中国力量：月之暗面的“Kimi”与深度求索的“R1”

月之暗面 & Kimi

Kimi K2 Thinking

惊人参数



1万亿 总参数 (MoE架构)



256k Token 超长上下文窗口



200-300次 连续工具调用能力

核心优势



极强的智能体工作流 (Agentic Workflows) 和信息整合能力。

深度求索 & R1

母公司：高飞 (High-Flyer) – 一家中国的对冲基金。

DeepSeek-R1 (推理模型)

市场影响



“斯普特尼克时刻”：其高性价比的开源模型给美国AI行业带来巨大震动。



2025年1月，其App在美国iOS商店的下载量一度超过ChatGPT。

核心理念



“开放权重 (Open-weight)”，鼓励开发者使用和创新。

全能选手与流量巨头：智谱的“GLM”和字节的“豆包”

智谱AI & GLM

GLM系列，特别是 GLM-4.7

核心技术

- “思考模式 (Thinking Mode)”
包含交错推理、保留推理和轮次推理，使复杂任务执行更稳定可控。

性能对标

- 在HLE基准测试中超越 GPT-5.1。
- 在 τ^2 -Bench交互工具调用基准上超越 Claude Sonnet 4.5。

擅长领域



- 智能体编程 (Agentic Coding)、前端UI和PPT生成。

字节跳动 & 豆包

豆包 (Doubao)

母公司：字节跳动

核心优势

- 依托于抖音等国民级应用，拥有巨大的用户基础和流量入口。
- 专注于为广大用户提供易于使用的AI助手服务。

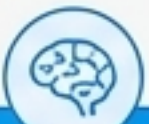
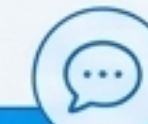
定位

- 面向消费市场的AI应用平台。

关系图谱全貌：AI世界的权力与连接



顶尖对决：谁是当前任务的最佳AI工具？

特性 (Feature)	 GPT-5.2 Thinking	 Claude Sonnet 4.5	 Kimi K2 Thinking	 GLM-4.7	 DeepSeek-R1
 母公司	• OpenAI	• Anthropic	• Moonshot AI	Zhipu AI	High-Flyer
 核心强项	• 综合智能、通用任务	• 软件工程、复杂代码	智能体 workflow、信息整合	• 智能体编程、办公自动化	• 推理、高性价比
 上下文窗口	N/A	200k / 1M	• 256k	200k	128k
 独特功能	• 可调思考时间	宪法AI、自主操作	透明推理、超多工具调用	• 可控思考模式	开放权重、MoE架构
 代表性基准	SWE-bench: 80%	编码基准领先	Agent基准表现突出	HLE > GPT-5.1	推理基准 > o1

“数据来源于公开的技术报告和第三方评测，旨在提供一个选型参考。”

结论：AI世界就像哆啦A梦的口袋，没有最好，只有最合适！

“大雄，你看，现在你明白了吧？
选择AI模型就像从我的口袋里挑选道具一样！”

如果你的任务是...

- 🔧 复杂编程和软件修复：
试试 Claude Sonnet 4.5 或 GPT-5.2!
- 🔍 需要自主完成的多步骤研究：
‘Kimi K2 Thinking 是个不知疲倦的好帮手!’
- 💡 追求极致的智慧和创造力：
‘GPT-5.2 Pro 或 Claude Opus 4.5 能给你惊喜。
- 💰 预算有限但需要强大能力：
‘DeepSeek 的开源模型性价比超高!’



“未来已来，选择正确的工具，你也能创造自己的奇迹！”