



AI大模型进化论

从“鹦鹉学舌”到“深度思考”

进化基石：Transformer 架构的诞生（2017）

从“逐字阅读”到“一目十行”

革命前 (RNN时代)



AI处理语言像小学生读课文，读了后面忘前面。

革命后 (Transformer)



引入“自注意力机制”，让AI能同时关注句子中所有词的关系，效率和理解力呈指数级上升。它是现代所有大模型（GPT, BERT, Llama）的共同祖先。

AI 的“iPhone 时刻”：ChatGPT 横空出世（2022）



1 亿用户 / 2 个月

通过“人类反馈强化学习”（RLHF）AI 终于能听懂人话，从复杂的工具变为流畅的对话伙伴。成为史上用户增长最快的应用。

Alibaba PuHuiTi 3.0 Heavy

群雄逐鹿之一：DeepSeek 的“极致性价比”革命



传统模型



DeepSeek(MoE)

混合专家模型 (MoE) – 每次只激活一小部分“专家”神经元来解决问题，而非调动整个大脑。

类比：如同一个专家团队，针对不同问题，只有相关领域的专家出来工作。

知识量：6710亿总参数
计算量：仅370亿参数参与计算

Alibaba PuHuiTi 3.0 Heavy

开源界的“价格屠夫”

凭借MoE架构，DeepSeek 实现了颠覆性的成本控制。

训练成本：约550万美元，
约为GPT-4 的 1/20。

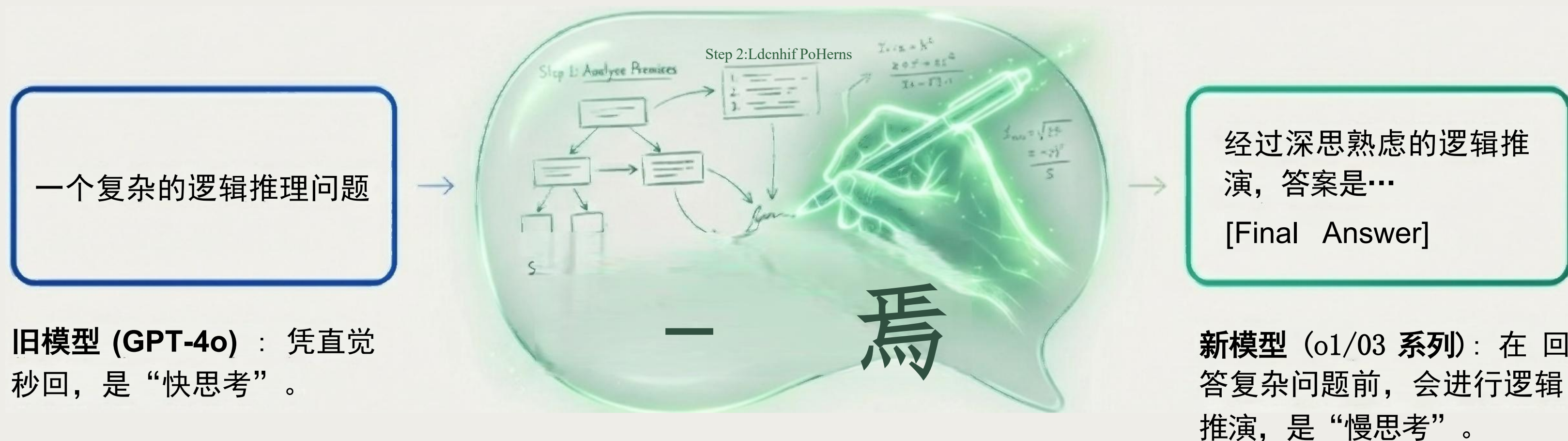


API价格：仅为OpenAI的
1/10~1/30。

Alibaba PuHuiTi 3.0 Heavy

群雄逐鹿之二：OpenAI 的“慢思考”进化

从快思考到慢思考 (System 1 → System 2)



成果：o3 模型在ARC-AGI（抽象推理测试）中得分高达 **88%**，超越人类平均水平。

群雄逐鹿之三：Google Gemini的原生多模态

天生就能同时理解文字、
图像、音频和视频，而非
后期拼凑。



能看懂长达数小时的视频，
听懂复杂的音频。

超长记忆：一次读完整个代码库

超长上下文窗口 (Context Window)

1 Million + Tokens = 一次读完7遍《哈利 · 波特》全集



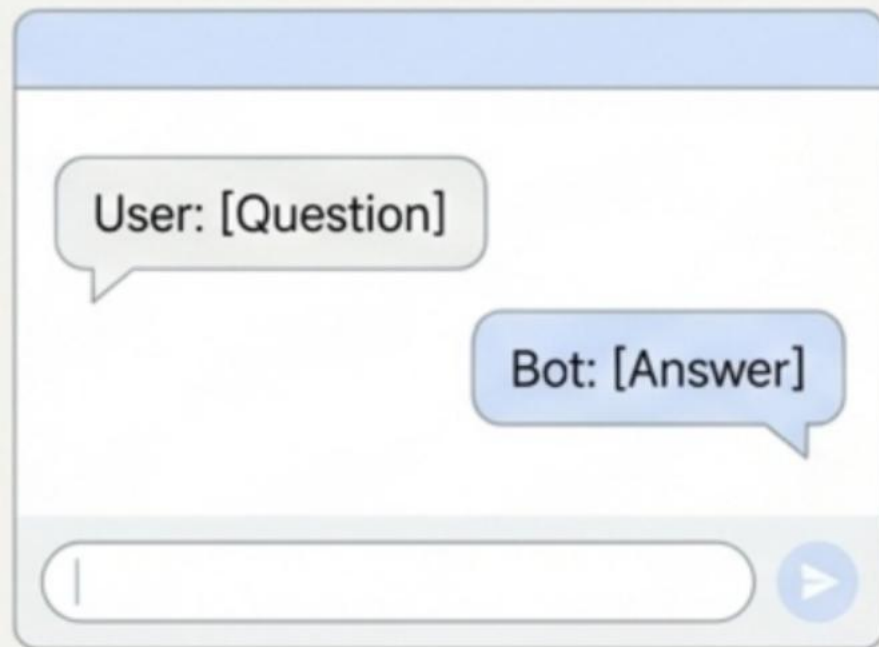
Gemini 3 Pro支持100万至200万 token。

你可以一次性把几百本书或整个项目的代码库扔给它。

它能全部记住并进行深度分析，拥有近乎无限的短期记忆。

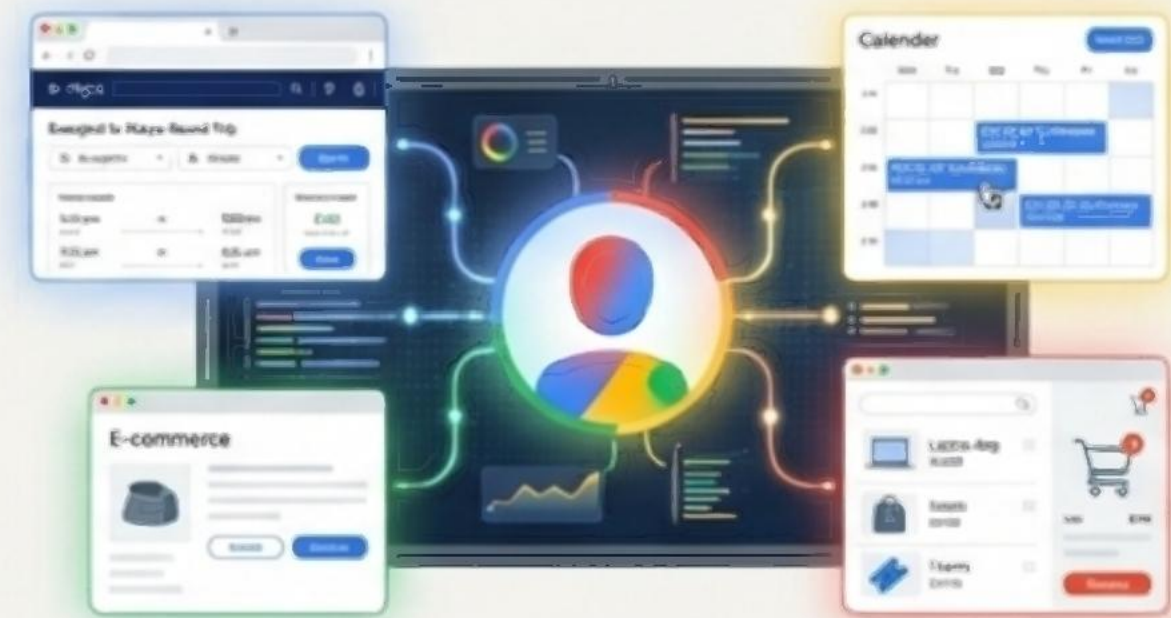
下一个物种：从“聊天工具”到“自主智能体” (Agent)

AGI（通用人工智能）指具备像人类一样，能解决从未见过问题的广泛认知能力。



过去 (Chatbot)

你问，它答。



未来 (Agent)

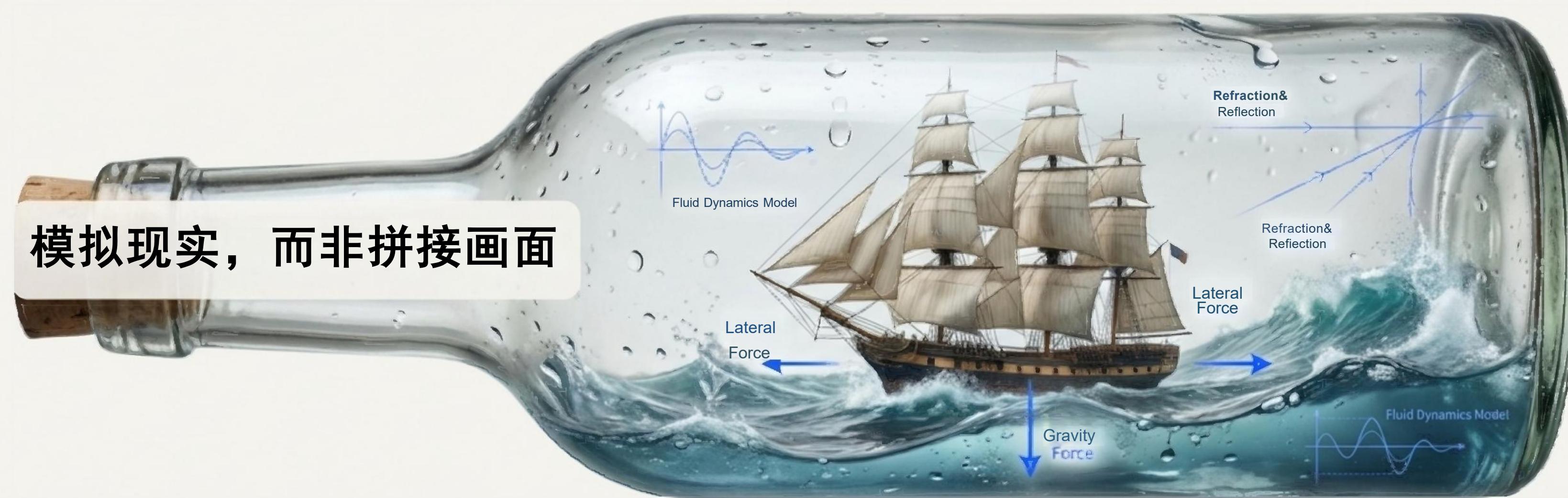
你给一个目标，它自主拆解任务、搜索信息、完成执行。

2025年被视为“自主时代”的开端。

睁眼看世界：世界模型（World Models）的崛起

核心突破：AI从理解“文字概率”进化到理解“物理规律”。

模拟现实，而非拼接画面



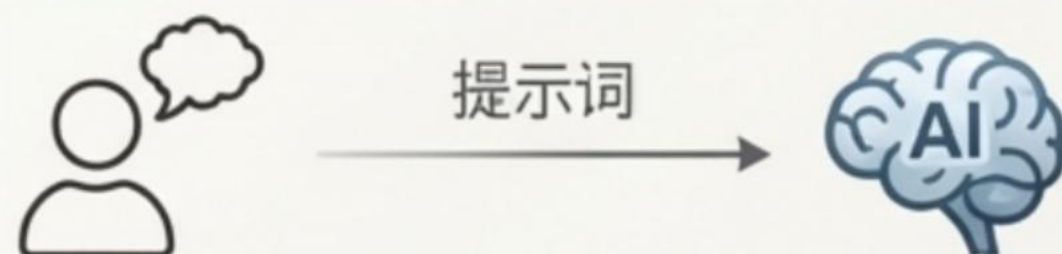
能力：懂得重力、碰撞、空间关系等现实世界的运作方式。

代表技术：OpenAI 的 Sora 2, Google 的 Gemini 3.0。

意义：这是AI从“纸上谈兵”走向“理解现实”的关键一步。

新的人机关系：从“提示词工程”到“上下文工程”

过去 (Prompt Engineering)



学习如何写好一句话指令。

未来 (Context Engineering)



学习如何管理AI的“记忆”和“背景信息”，让它更懂你的长期需求。

AI的局限：

- 幻觉 (Hallucinations): AI 仍会一本正经地胡说八道。
- 偏见 (Bias): AI 会继承训练数据中的社会偏见。

我们的新角色：AI 时代的“指挥官”与“审核者”

Human-in-the-loop （人在回路中）



AI的计算力

不知疲倦的“执行者”，负责
计算与操作。

人类的判断力与创造力

拥有判断力与创造力的“指挥官”，
负责设定目标与最终决策。

未来已来，你准备好了吗？

AI的发展速度是指数级的（从2020到2025仅仅5年）。
拥抱变化，掌握与AI协作的能力，是新时代的核心竞争力。

