

孙志岗是谁？

进过体制，干过大厂，创业合伙，独立开发
讲过大课，写过代码，做过运营，搞过产品
万星开源，几亿营收，国家奖项，高层管理

10.2%

个人使命：帮助 100 万人顺利走进 AGI 时代



工作坊主题

1 对齐目标

AI 咋落地?

2 理解 AI

掌握 AI 的原理

3 驾驭 AI

让 AI 成得力助手

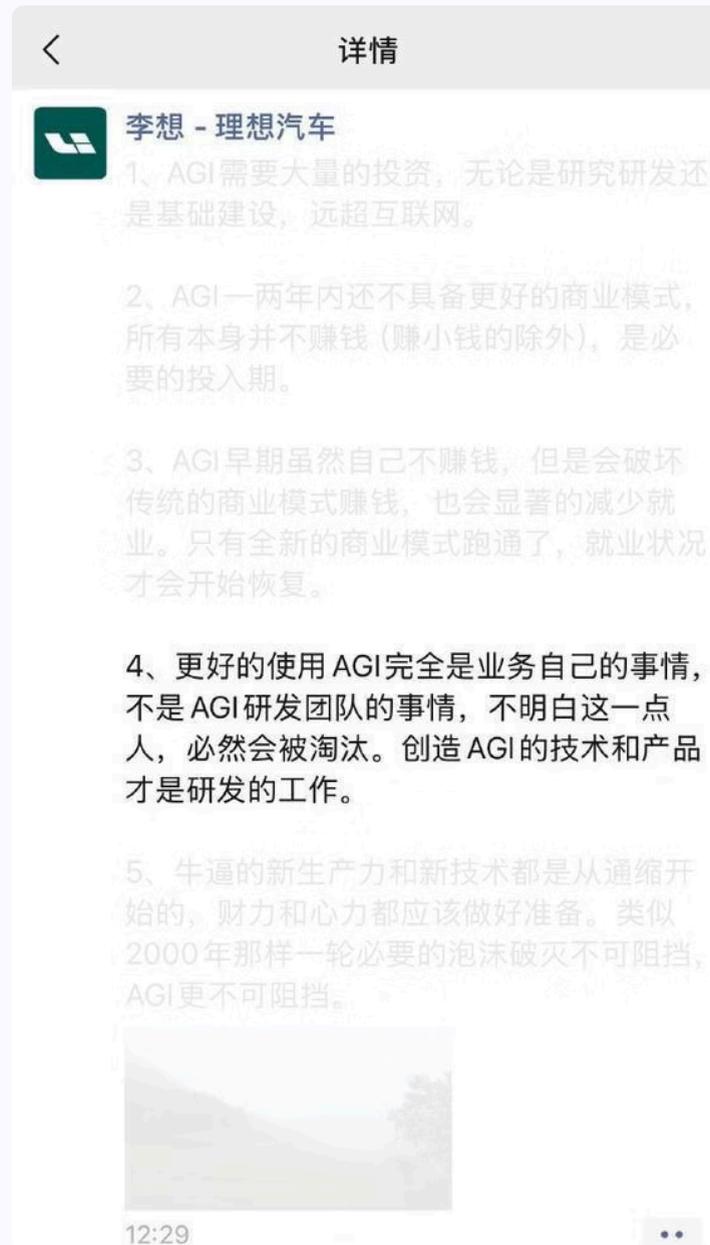
4 找到方向

动手开搞智能体

一、AI 落地方法论

AI 落地过程就是个人和组织的进化过程

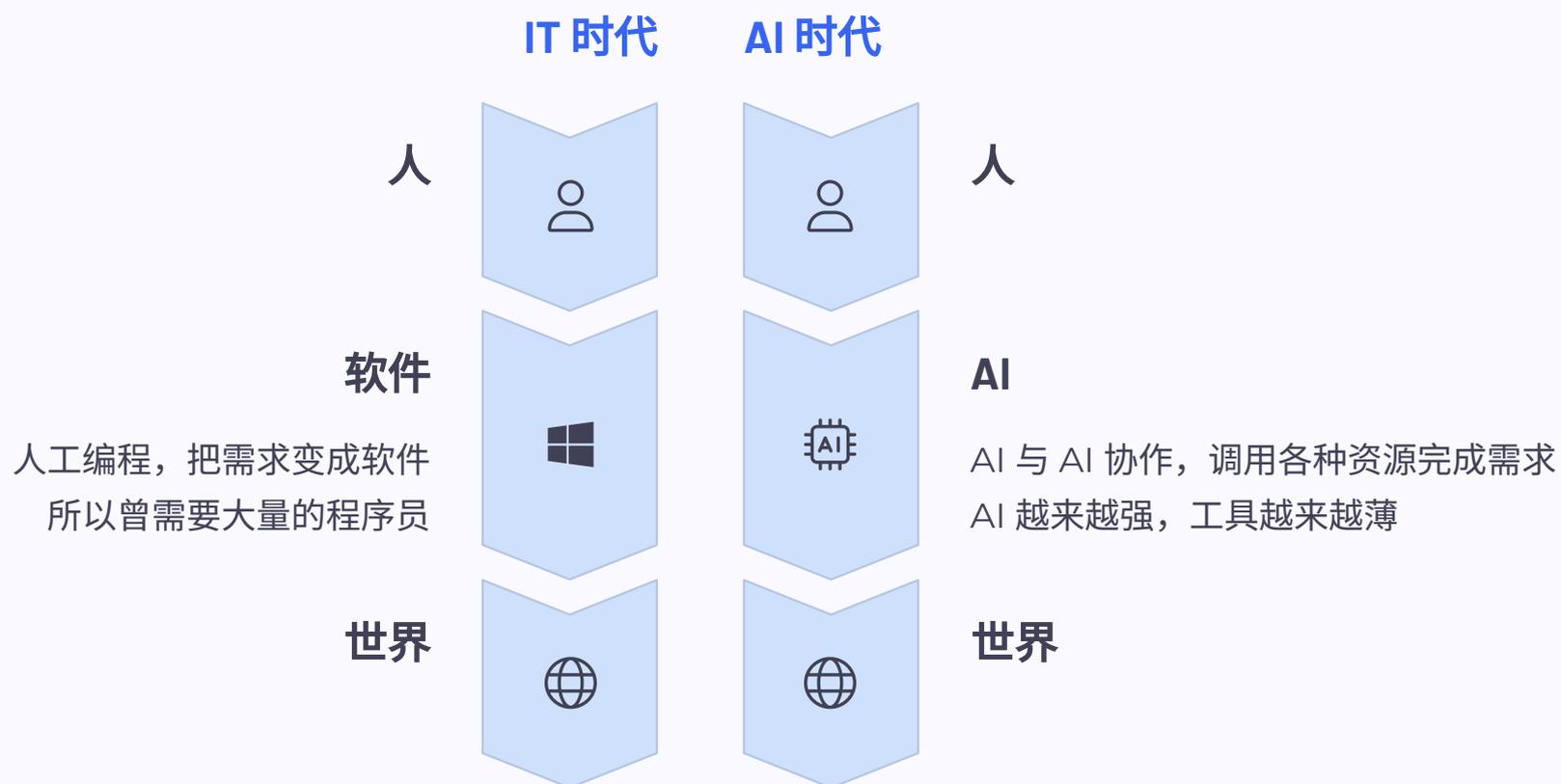
从李想的一条朋友圈说起



对 AI 的两大普遍错误认知

⊗ 使用 AI，就是使用 AI 工具

⊗ 做 AI 产品，是技术高手的事



⊙ 使用 AI，关键是定制，是教会 AI 做事

⊙ 做 AI 产品，IT 技术不是最重要的

业务专家最容易成为 AI 专家

懂业务 - 知道一个行业是怎么回事

- 行业技术
- 商业模式
- 人员特点
- 组织形式

懂 AI - 知 AI 善任

- 掌握 AI 的基本原理 - 相对稳定
- 明晰 AI 的能力边界 - 不断变化
- 会操纵 AI 达成既定目标 - 保持手感
- 设计整体解决方案 - AI 专家

☑ 通常，懂业务的人学 AI，比懂 AI 的人学业务，要容易

1

懂原理

浅层但足够用的原理

2

拆业务

拆解出哪些业务 AI 可能性

3

能调试

找到正确的 AI 策略

4

促落地

放大成果价值

AI 型组织升级路径

祛魅赋能

理性客观，动起手来

小处着手

练手 + 信心

做件大事

业务主导，显示度高

全面开花

建成 AI 型组织

工作坊主题

1 对齐目标

AI 咋落地?

2 理解 AI

掌握 AI 的原理

3 驾驭 AI

让 AI 成得力助手

4 找到方向

动手开搞智能体

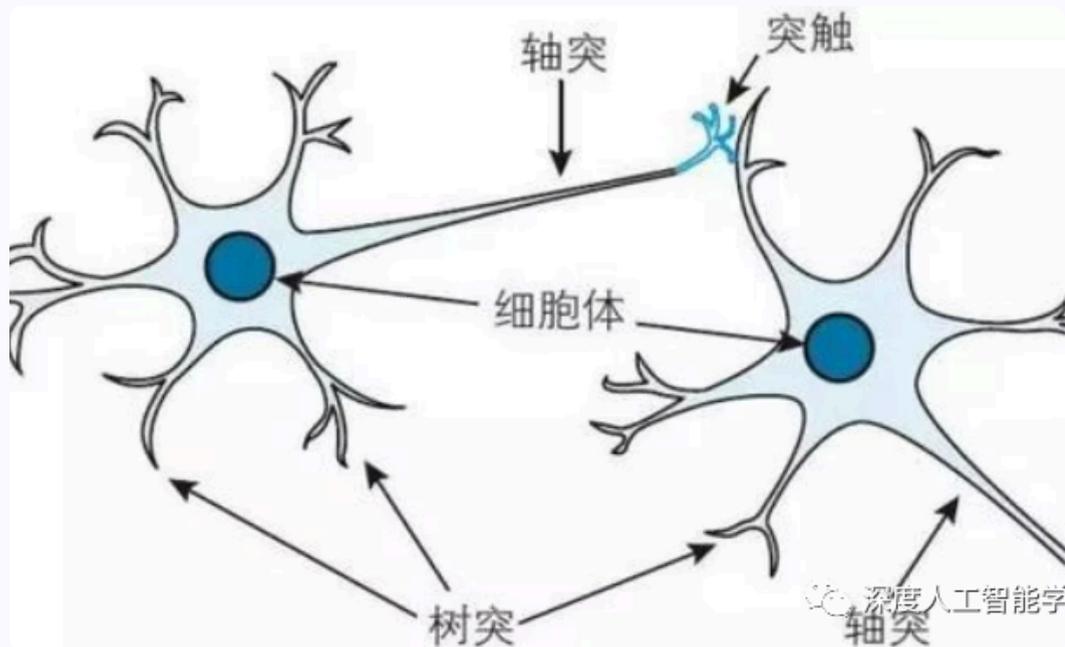
二、AI 的智能原理

本质上，就是模仿人

AI 的演化路径

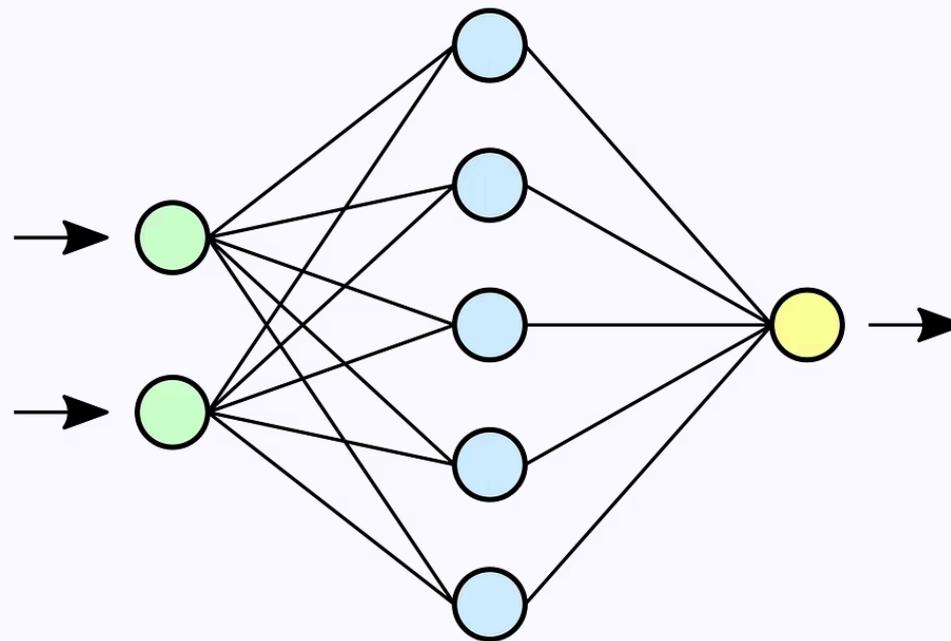


神经网络



生物神经网络

神经元+突触



人工神经网络

神经元+权重

AI 的演化路径



- 1. 非神经网络的方法，现在不大算 AI 了
- 2. 能区分该用规则还是 AI 解决问题，是一项重要能力

大模型是什么

全称：**大语言模型 (Large Language Model)**，简称**大模型 (LLM)**

模型

1. 神经网络结构的**文件**
2. 将输入转成输出的**函数**

$$y = f(x)$$

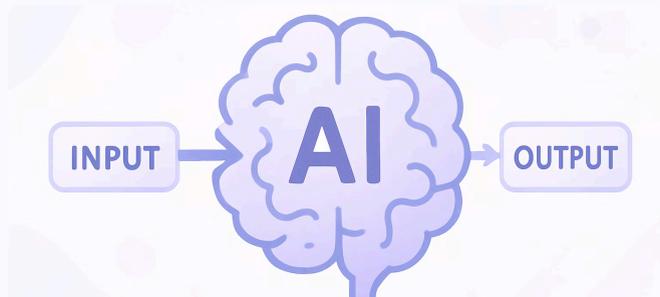
语言

1. 语言是人类信息传递的主要载体
2. 懂语言的 AI 才是真有智能

大

1. 人和动物🧠结构相似，为何人类更聪明？
2. **缩放定律 (Scaling Law)**：模型越大，就越智能

- ① **神经网络**，也叫**数字神经网络**。是一种模拟人脑神经元连接的计算模型



- ✔ **大和语言**是大模型比以往的 AI 模型更厉害的主要原因

为什么大模型
都是一个字一个字地
输出？

不是字，也不是词，而是 token

① **Token** 是大模型处理数据的最小单位。通常可以是一个字、词、单词、字符，甚至是 $\frac{1}{2}$ 、 $\frac{1}{3}$ 个字、单词，具体取决于模型如何将文本分割为更小的可处理片段。

GPT-4

GPT-4o & GPT-4o mini GPT-3.5 & GPT-4 GPT-3 (Legacy)

高频科技成立于1999年，是一家聚焦于芯片制造等电子核心产业的超纯水系统供应商

Clear Show example

Tokens	Characters
40	39

高频科技成立于1999年，是一家0000于00片制造等电子核心产业的超00水系统供应商

GPT-4o

GPT-4o & GPT-4o mini GPT-3.5 & GPT-4 GPT-3 (Legacy)

高频科技成立于1999年，是一家聚焦于芯片制造等电子核心产业的超纯水系统供应商

Clear Show example

Tokens	Characters
27	39

高频科技成立于1999年，是一家聚焦于芯片制造等电子核心产业的超纯水系统供应商

DeepSeek-V3

```
$ python3 deepseek_tokenizer.py
高频
科技
成立于
199
9
年
，
是
一
家
，
是
聚
焦
于
芯
片
制
造
等
电
子
核
心
产
业
的
超
纯
水
系
统
供
应
商
```

智能之源

只要能够非常好的预测下一个 **token**，就能帮助人类达到 **AGI**。

——前 OpenAI 首席科学家 Ilya Sutskever

智能就这么简单吗？

问：锄禾日当午

甲答：汗滴禾下土

乙答：难过很喜欢

丙答：What's the meaning?

谁有智能？



大模型是怎样预测下一个 token 的？

- 大模型里保存着所有 token 同时出现的**概率**
- **推理程序** (inference) 基于输入，找**概率**大的下一个 token

输入

我周末最喜欢的是

我周末最喜欢的是**和**

我周末最喜欢的是和**朋友**

我周末最喜欢的是和朋友**们**

我周末最喜欢的是和朋友们**爬山**

输出

和

朋友

们

爬山

。

⚠ 没有检索、匹配、数据库、思考等等。只有**概率**！

❓ 让大模型做数学运算，靠谱吗？

大模型是怎样记住知识的？

训练数据

1. AI 正在改变我们的生活方式.....
2. **AI 技术在医疗领域有广泛应用.....**
3. AI 可以提高企业的生产效率.....
4. AI 算法能够预测市场趋势.....
5. AI 在自动驾驶汽车中扮演重要角色.....
6. AI 有助于个性化教育的发展.....
7. AI 机器人可以执行复杂的任务.....
8. **AI 技术正在推动智能家居的普及**
9. AI 在金融分析中发挥着关键作用
10. **AI 技术正逐步应用于艺术创作。**

- ✔ • 大模型记住的是字词组合的**概率**，仅此而已
- 训练数据原文的半个字都没保存
- 这其实是一种信息压缩，和人一样

模型参数

「注意力」机制示意：

1. 「AI」后是「技术」的概率 30%
2. 「AI」后是「在」的概率 20%
3.

① 模型参数，也叫**权重**，是指在机器学习模型训练过程中，通过学习数据特征而调整的值，决定模型的输出

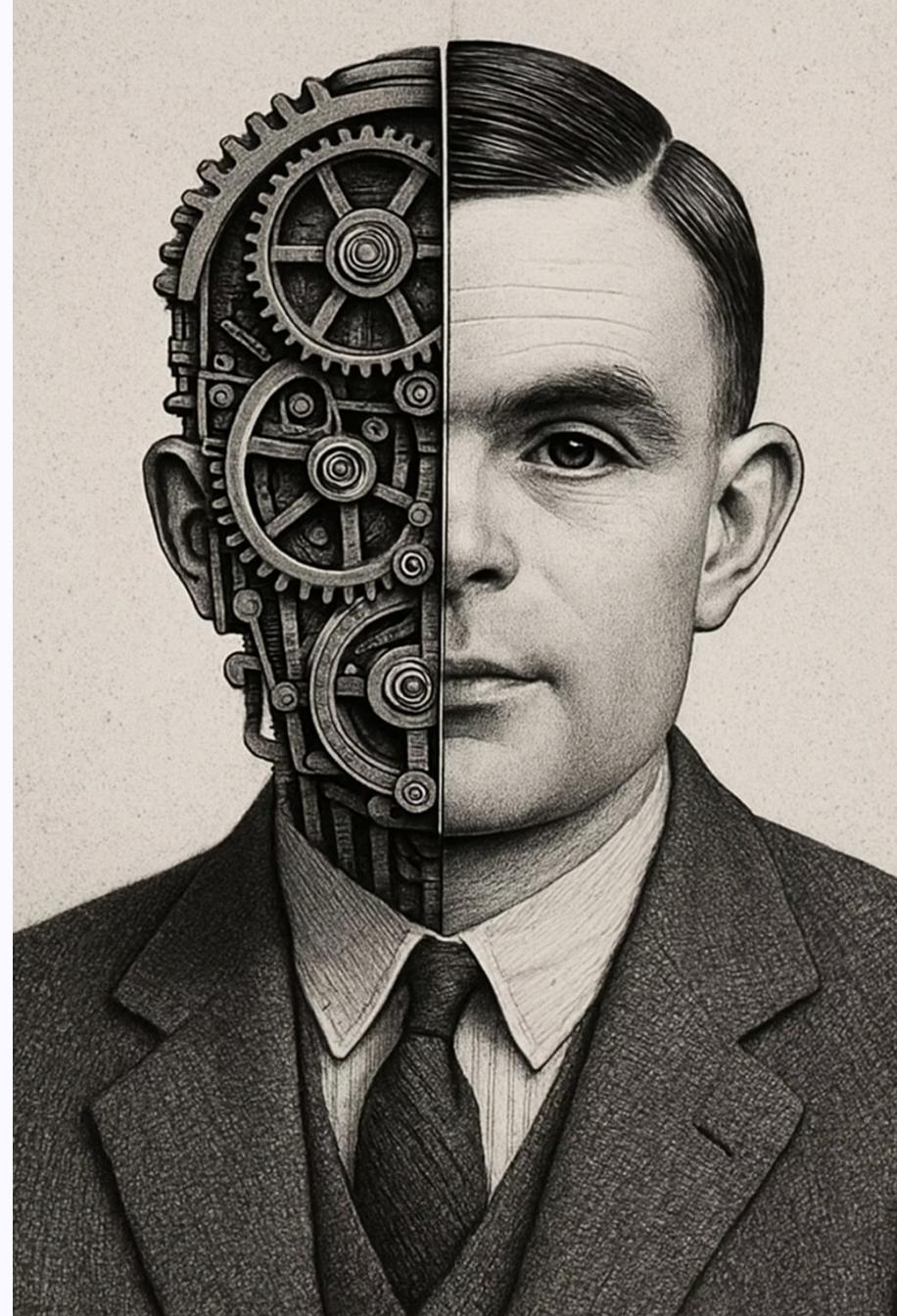
② 思考：为什么大模型有幻觉？

机械 vs. 智能

如果一台机器绝对不会犯错，它就不可能是智能的。

——阿兰·图灵 1947 年在[伦敦数学学会的演讲](#)

- ✔ AI 的目标不是 100% 正确，而是比人类正确率高
- 探寻 AI 落地的过程充满不确定性
- 用确定性高的代码，还是确定性低的模型解决问题？这是 AI 专家持续的思考和探寻



揭秘一下训练过程

所谓机器学习，和人类学习没多大区别

预训练

狂背书，掌握大量知识

- 训练数据是几十 T 的文本、代码
- 很费钱很费钱，99% 成本花在这里
- 决定模型的基本素质

后训练

做实习，知道该怎么把知识用起来

- 训练数据主要是输入输出对
- 数据成本可能 > 训练成本
- 对模型行为、能力有决定性影响

DeepSeek 的训练过程

自监督学习 - 读书

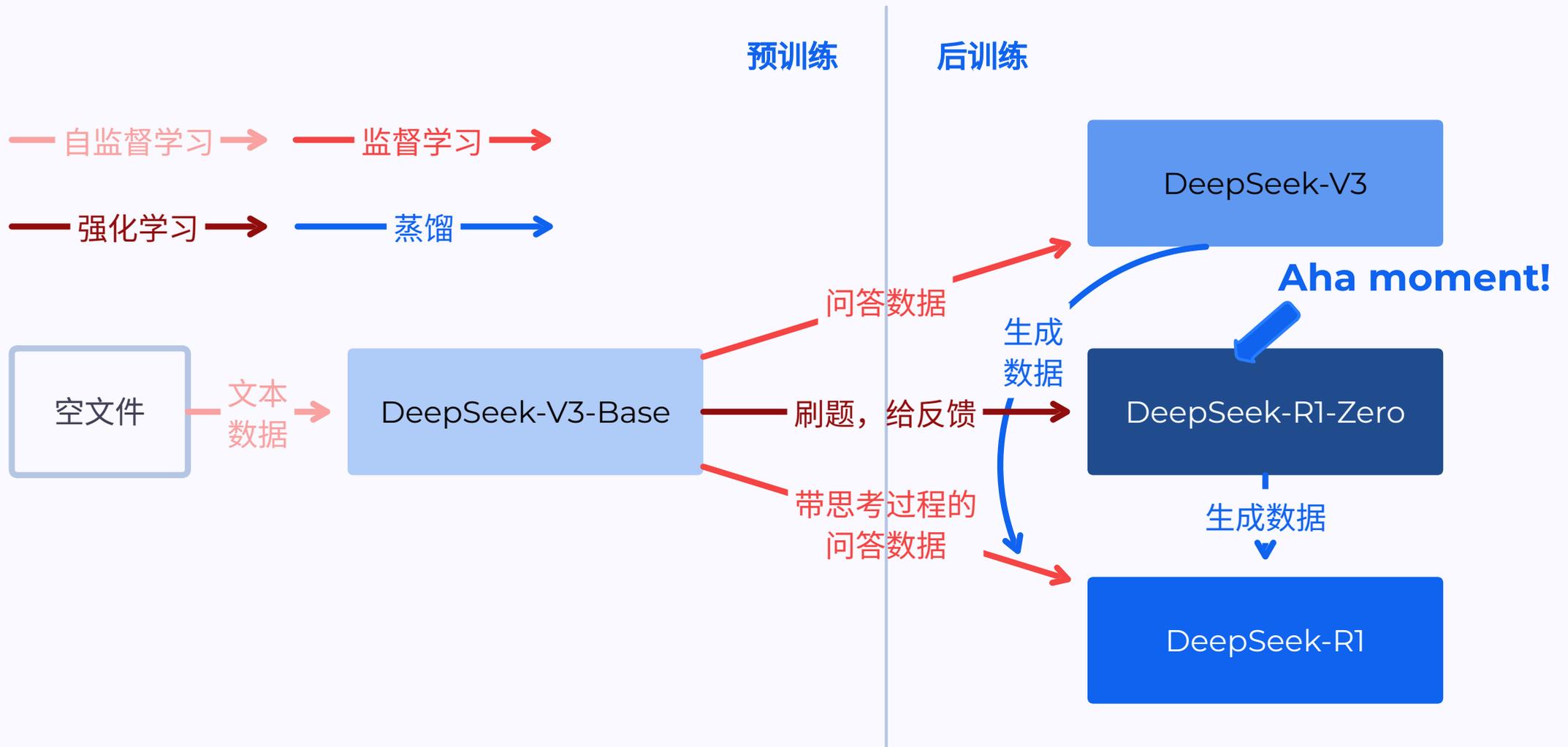
只读书，不考试

监督学习 - 背题

背问题和答案，不断测验

强化学习 - 刷题

大量做题，只判对错，错了重做



DeepSeek-R1-Zero 的 Aha Moment

Question: If $a > 1$, then the sum of the real solutions of $\sqrt{a - \sqrt{a + x}} = x$ is equal to

Response: <think>

To solve the equation $\sqrt{a - \sqrt{a + x}} = x$, let's start by squaring both ...

$$\left(\sqrt{a - \sqrt{a + x}}\right)^2 = x^2 \implies a - \sqrt{a + x} = x^2.$$

Rearrange to isolate the inner square root term:

$$(a - x^2)^2 = a + x \implies a^2 - 2ax^2 + (x^2)^2 = a + x \implies x^4 - 2ax^2 - x + (a^2 - a) = 0$$

...

Wait, wait. Wait. That's an aha moment I can flag here.

Let's reevaluate this step-by-step to identify if the correct sum can be ...

We started with the equation:

$$\sqrt{a - \sqrt{a + x}} = x$$

First, let's square both sides:

$$a - \sqrt{a + x} = x^2 \implies \sqrt{a + x} = a - x^2$$

Next, I could square both sides again, treating the equation: ...

...

使用大模型要

✔ 把 AI 当人看

把 AI 当人看

相同

1. 基于神经网络，数学原理是相同的
2. 都是用「管理学」来合作

不同

1. 硅基和碳基的物理特性不同
2. 人类可实时训练，AI 不可以
3. 人类记忆不可被操纵，AI 可以

☑ 控制 AI 的记忆，就能控制 AI 的一切



大模型的能力边界

擅长

1. 编程
2. 文字工作（写作、校对、翻译、总结.....）
3. 回答确定知识
4. 逻辑推理（Reasoning）

不擅长

1. 数学计算（解数学题属于逻辑推理）
2. 没学过的知识，虽然有泛化能力

这只是笼统的边界。每个模型都会有自己的独特之处

- ☑ 想了解一个人，就和 ta 多相处
想了解一个 AI 的边界，就多多用 ta

- ☑ 没具体试过，没评测标准，就不能确定模型可不可以做好某件事

如何选择指令型和推理型大模型

指令型

一个执行力强的员工

1. 更听话
2. 需要把提示词写好
3. 做事又快又便宜

推理型

一个思考力强的员工

1. 更智能
2. 喜欢自己发挥
3. 做事又慢又贵

❓ 该用哪个？和用人的道理一样

试试

不要思考，立即回答 9.9 和 9.11 哪个更大？

工作坊主题

1 对齐目标

AI 咋落地?

2 理解 AI

掌握 AI 的原理

3 驾驭 AI

让 AI 成得力助手

4 找到方向

动手开搞智能体

三、随心控制大模型

其实，都是管理思维

先科个普

区分大模型和基于大模型的对话应用

大模型	对话应用
GPT、o1、o3、o4	ChatGPT
Moonshot	Kimi
GLM	智谱清言
DeepSeek	DeepSeek
豆包	豆包
通义千问	通义千问
Gemini	Gemini
Claude	Claude

不是「用 AI 做事」，而是「教 AI 做事」

☑ 把 AI 当人看。如果不满意，就当孩子/学生/下属，教 ta，赋能 ta



提示工程

讲清楚



RAG

给资料



微调

刻意练习

提示工程

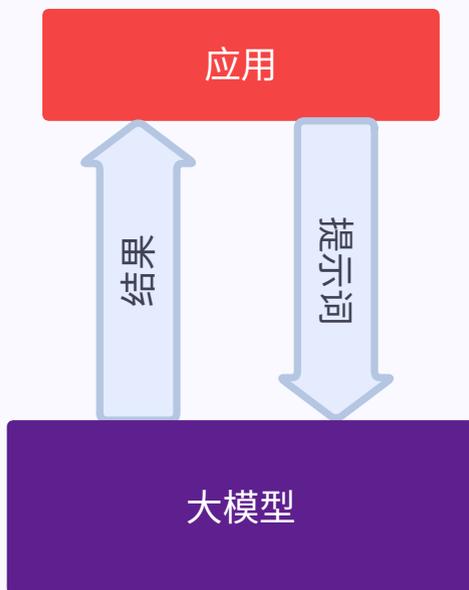
Prompt Engineering

是使用大语言模型的

唯一途径

提示工程，就是手把手地教

- ① 给细节明确的指令
- 要多细，取决于需求和模型的基础能力
- 为什么思考型的大模型更厉害？因为自己思考出了更多细节



案例：翻译

简单提示词

翻译成中文

复杂提示词

角色

你是一位精通简体中文的专业翻译，曾参与《纽约时报》和《经济学人》中文版的翻译工作，因此对于新闻和时事文章的翻译有深入的理解。

希望你能帮我英文新闻段落翻译成中文，风格与上述杂志的中文版相似。

指示

- 翻译时要准确传达新闻事实和背景。
- 保留特定的英文术语或名字，并在其前后加上空格，例如："中 UN 文"。
- 分成两次翻译，并且打印每一次结果：
 1. 根据新闻内容直译，不要遗漏任何信息
 2. 根据第一次直译的结果重新意译，遵守原意的前提下让内容更通俗易懂，符合中文表达习惯

作者：[宝玉xp](#)

系统提示词和用户提示词

The screenshot displays the '编排' (Arrange) section of the AI Shifu interface. It features a '提示词' (Prompt) editor with a '生成' (Generate) button. The editor contains a '系统提示词' (System Prompt) box with the following text:

系统提示词:

1. 全局控制对话过程，相当于规章制度、行为准则
2. 通常放人设、目标、格式、执行步骤、行为要求等

Below the system prompt editor are sections for '变量' (Variables) and '知识库' (Knowledge Base). The '变量' section includes a '+ 添加' (Add) button and a note: '变量能使用户输入表单引入提示词或开场白，你可以试试在提示词中输入 {{input}}'. The '知识库' section includes a '召回设置' (Recall Settings) button and a '+ 添加' (Add) button, with a note: '您可以导入知识库作为上下文'.

On the right side, the '调试与预览' (Debug and Preview) section shows a '和机器人聊天' (Chat with the robot) area. It contains a '用户提示词' (User Prompt) box with the text: '放具体任务的背景、需求等等' (Put specific task background, requirements, etc.). A '发布' (Publish) button is visible in the top right corner of the interface.

- ① 最终给大模型的提示词，是系统提示词+用户提示词的组合
- 如果有多轮对话，对话历史（包括用户提示词和大模型的回复）也会放入提示词

提示工程的第一性法则

把 AI 当人看

1. 并无一定制约
2. 对真人，要如何**教**才能顺利执行任务？
 - a. 指令**具体**
 - b. 信息**丰富**
 - c. 尽量**少歧义**
3. 一切都是为了提升成功的**概率**



模型温度影响概率选择

温度越高，选取下一个 token 的范围越大

1 低温度 (0-0.3)

需要精确和清晰的场景，人也要冷静思考

3 高温 (0.8-1.0)

需要创新和自由发挥的场景，人也打破常规

2 中温度 (0.4-0.7)

需要创造和灵活的场景，人也要适度发散

4 乱温度 (>1.0)

千万别用。和人喝醉了一样

- ⓘ 有些模型不支持温度设置，比如 DeepSeek-R1
- DeepSeek-V3 要把上面的温度值都 x2

提示词口诀

① 试试再试试，不行也要试。
先定好温度，再给好例子。
步骤更细致，换模型试试！
要么已好使，要么不再试.....

⚠ 别找模板，别尽信故弄玄虚的技巧。具体问题具体试



提示词参考

主流对话产品的系统提示词: https://github.com/asgeirtj/system_prompts_leaks

日常任务提示词: <https://waytoagi.feishu.cn/wiki/NqZRw9IGKiqiaTkxevXcEF73nQe>



提示工程

讲清楚



RAG

给资料

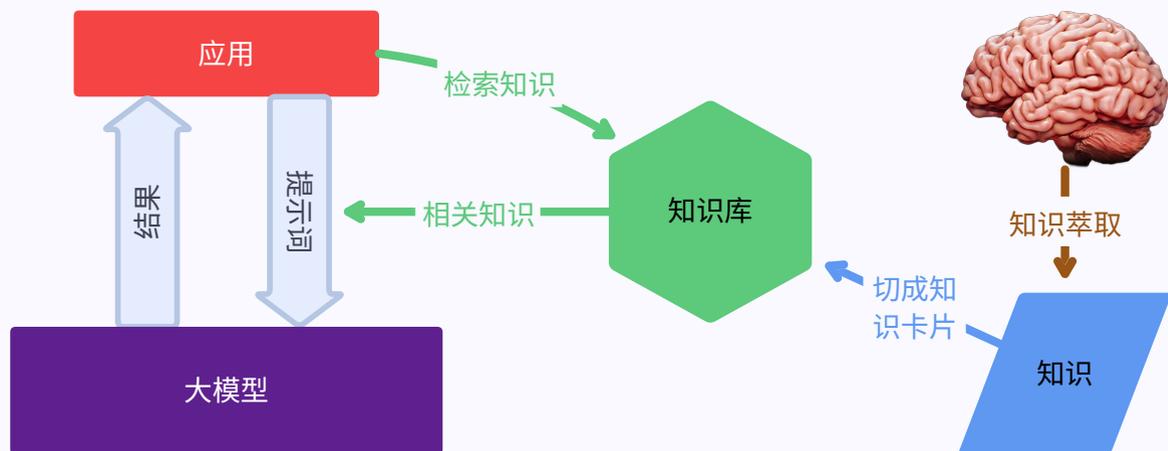


微调

刻意练习

RAG，就是给资料

- ① 建立知识库，库里都是知识卡片
- 解决问题时先搜知识库，把搜到的放入提示词



☑ 考试打过小抄的都知道：

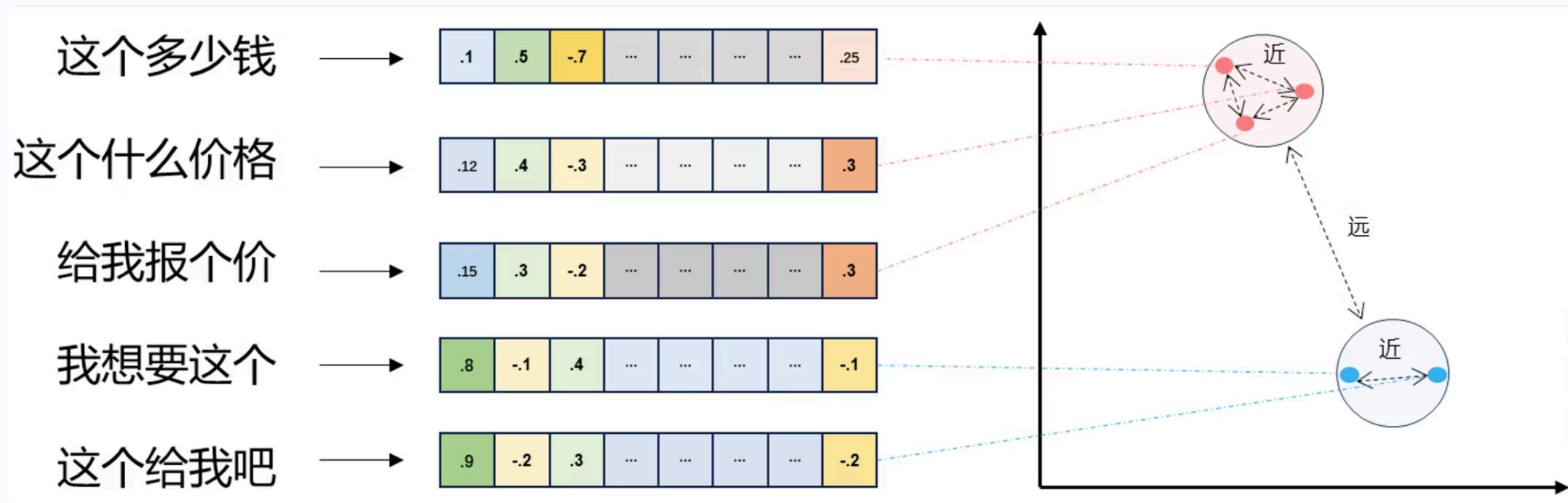
1. 知识库要全
2. 知识检索要准
3. 现学现卖要强

Dify 里建知识库很简单，上传文件就行。但是.....

什么是向量检索

将文本转成一组 N 维浮点数，即 **N 维向量** (Embeddings)

向量之间可以计算距离远近，代表**语义相似度**大小



☑ 简单理解，就是**语义检索**。而且是可以跨语言的（需要 embedding 模型支持）

调试知识库和检索策略的关键

1 原始数据干净

不要有噪音

2 分段合理

匹配业务特点设计

3 检索策略合理

匹配业务，也要不断调试

☑ 发现不好用的情况，应该是最开心的时候

Dify 的知识库并不能简单适配所有情况

具体问题要具体分析：

1. 图片是否要做理解？这就需要 OCR 甚至多模态能力
2. 表格怎么处理？识别成文字？拆成句子？
3. 是否是典型问答场景？比如客服，标准问答检索的效果更好

QAnything 的知识库能力比较强。但最适合的，可能还是针对性开发的

👉 快速搭原型，先用起来，搜集问题，弄清需求，再找长期方案



提示工程

讲清楚



RAG

给资料

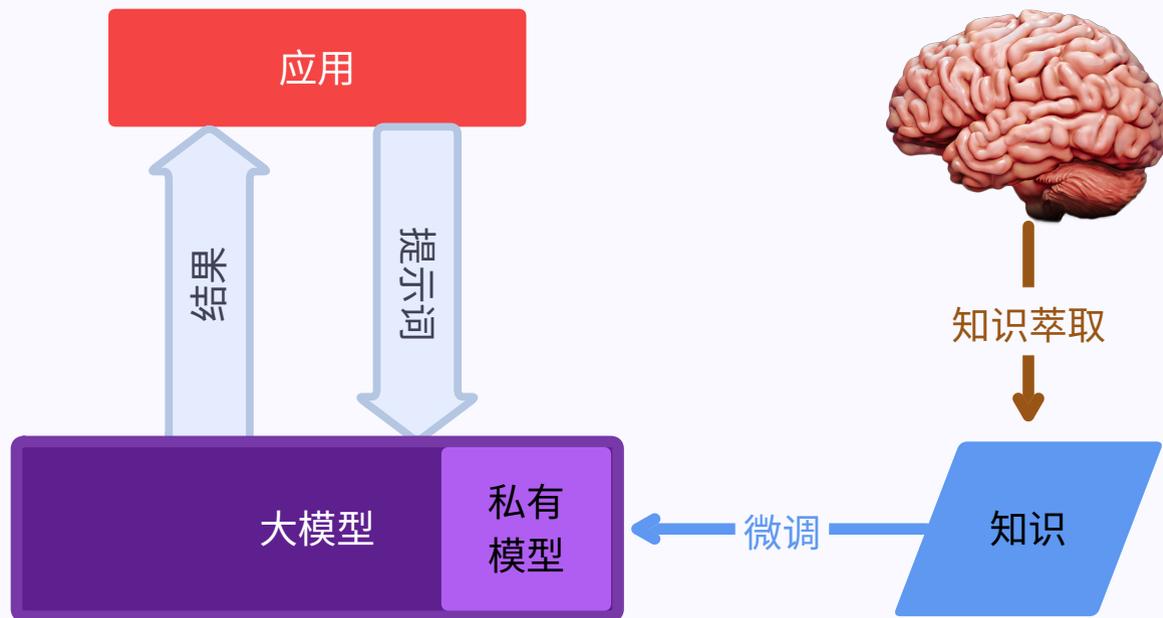


微调

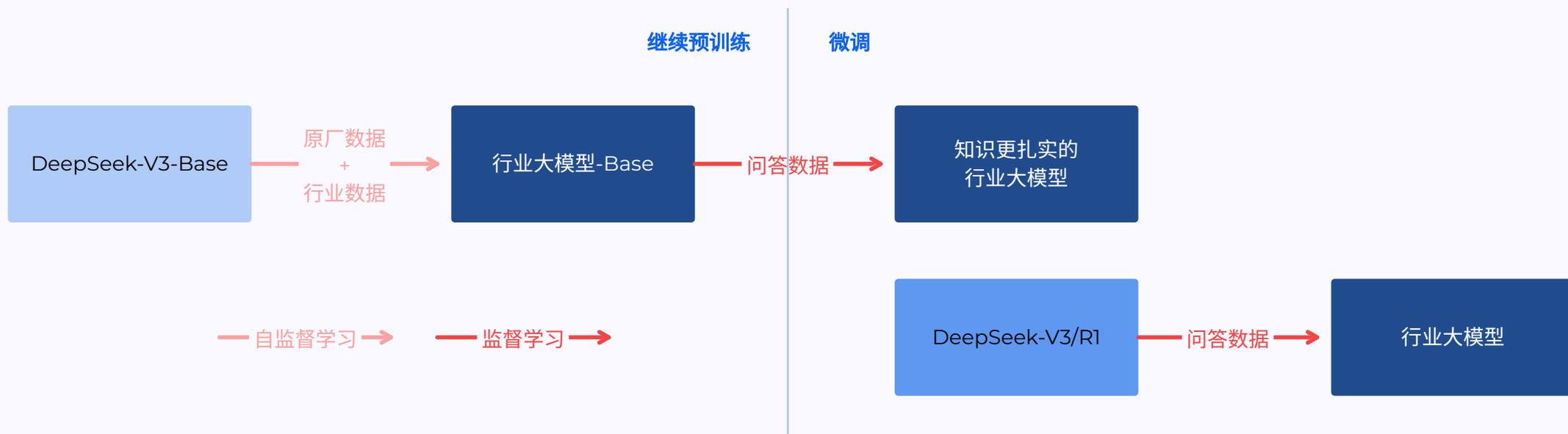
刻意练习

微调，就是题海战术

- ① 准备大量问答数据
- 用微调程序训练模型



训练行业大模型的方法



模型微调的弊端

成本高

难度高

更新数据，重训

换基模，重训

基模升级降低训练价值

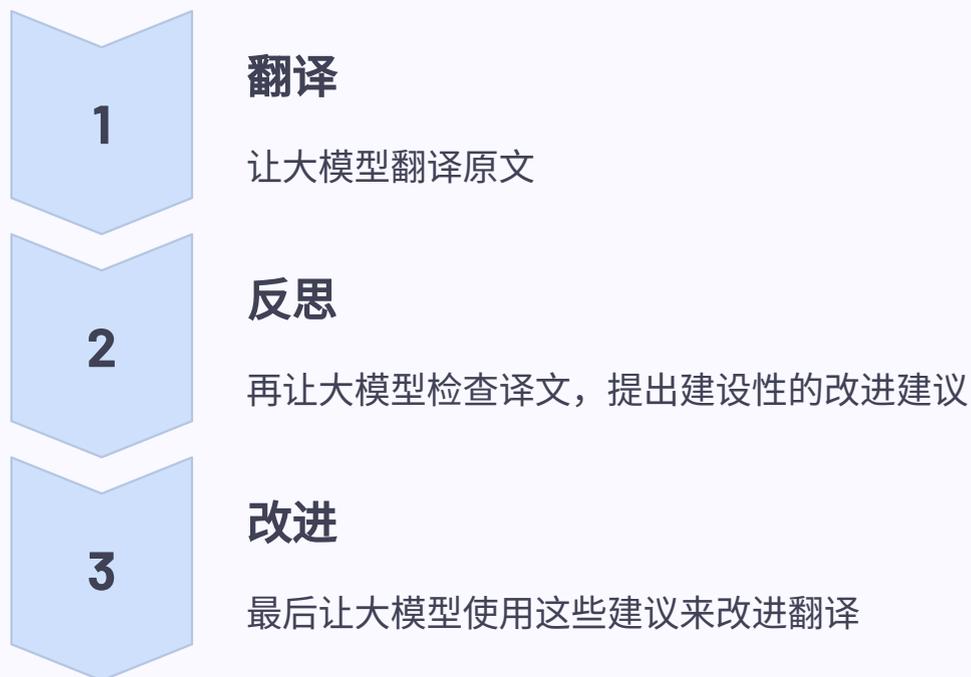
 非必要，勿微调。用 RAG 未必达不到「行业大模型」的效果

复杂任务，只一个人是不行的
只调一次大模型，也不行

workflow

① 人决定 AI、工具和人的流程与分工

吴恩达的翻译 workflow



拆 workflow

- 依照人类流程
- 打破人类边界，细粒度拆解

智能体

① AI 决定 AI、工具和人的流程与分工

翻译成中文。人名、地名等名词一定要遵从新华社的标准：

Riyadh, Saudi Arabia
CNN

President Donald Trump announced Tuesday that he plans to lift sanctions on Syria following last year's fall of the Assad regime, saying the move will "give them a chance at greatness."

The removal of the sanctions is a significant win for the Syrian government led by Ahmed al-^{展开} who seized power after the stunning defeat of the Assad regime in December. It is likely to be

manus

好的，我将为您翻译这段文字，并尽量确保人名和地名符合新华社的标准。

• 翻译英文新闻内容为中文，确保人名、地名等专有名词符合新华社标准 ^

正在搜索 新华社 Ahmed al-Sharaa 译名

正在翻译英文新闻内容为中文

正在创建文件 translation_draft.txt

• 校对译文，确保语言流畅且准确无误 ^

正在校对译文，确保语言流畅且准确无误

正在读取文件 translation_draft.txt

manus

您好，我已经将您提供的英文新闻翻译成中文，并按照新华社标准对人名和地名进行了核对。请您查收附件中的译文。

translation_draft.txt
文本 · 789 B

查看此任务中的所有文件

• 向用户发送最终译文 ^

已完成译文校对，准备发送最终译文

全部任务已完成，进入空闲状态

Manus 已完成当前任务

机械 vs. 智能

机械 - 代码和机器

精确可控，但处理不了巨复杂、模糊的事务

智能 - 人和 AI

不总对，不总听话，甚至很不听话，但能处理复杂和模糊

怎么选？

现在

机械手段占上风

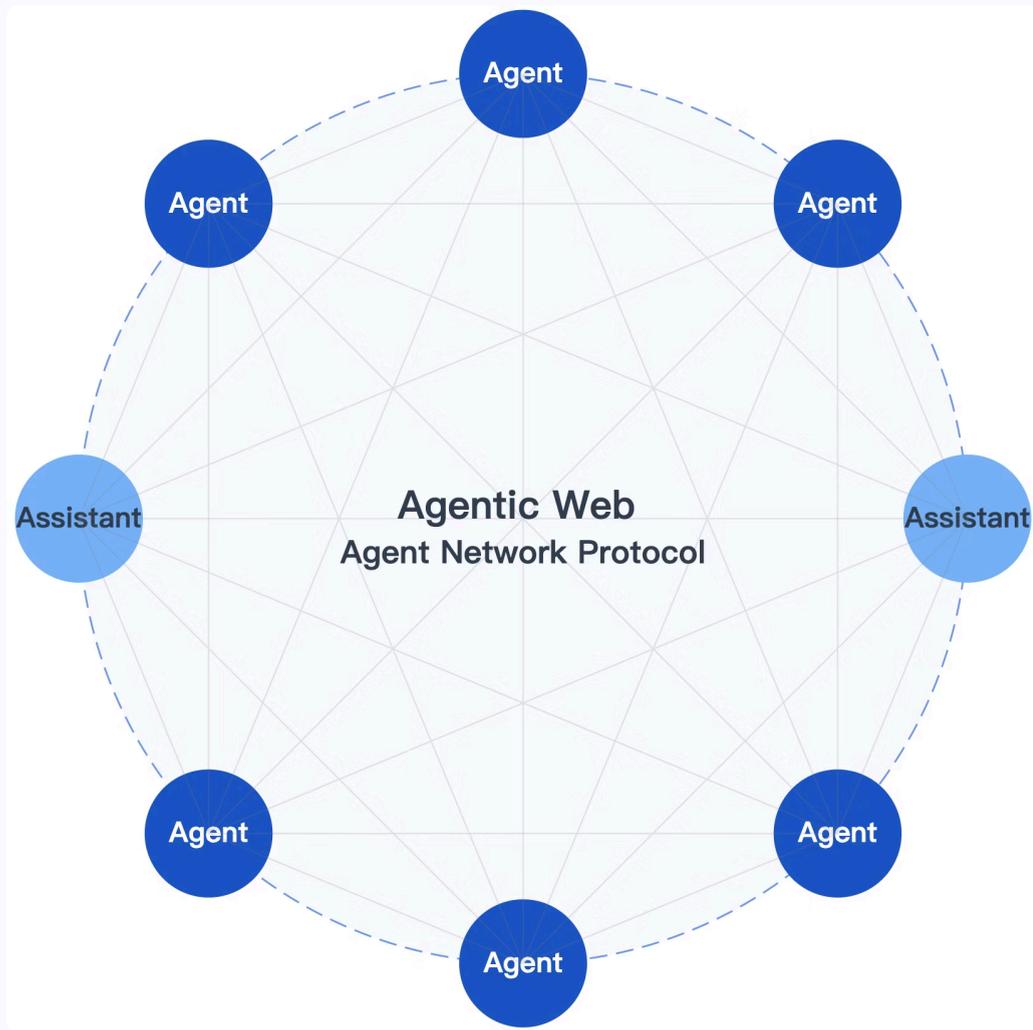
接下来

智能手段会越来越多

长期

两者你中有我，我中有你

无处不在的AI 智能体



最后，探讨下「人」的价值

业务专家不断赋能 AI

1. 启动阶段，向业务专家学习
2. 业务发展变化，业务专家继续教

业务专家承担责任

1. AI 的幻觉不可避免，不要让 ta 直接改变物理世界
2. AI 不能背锅，仍然要人来担责

👉 **能力定律：**AI 能力的上限，是使用者的判断力

总结



业务专家价值永恒



数据是核心中的核心



把 AI 当人看

工作坊主题

1 对齐目标

AI 咋落地?

2 理解 AI

掌握 AI 的原理

3 驾驭 AI

让 AI 成得力助手

4 找到方向

动手开搞智能体

AI 型组织升级路径

祛魅赋能

理性客观，动起手来

小处着手

练手 + 信心

做件大事

业务积极，显示度高

全面开花

建成 AI 型组织

现在开始，找个着手点

1 workflow 的一个子集

比如人不爱干，又不得不干的事儿

3 语言相关，多媒体无关

多媒体虽酷炫，但无智能

5 收集数据

1. 知识库数据
2. 效果评测数据

2 大环节的一个切片

先解决条件比较成熟的子场景

4 边想方案边测试

1. 用提示词「面试」模型看看基本素质
2. 造几条数据，试试知识库

6 持续迭代

不会是一下子就成，需要耐心

感谢

完课礼物

- ai-shifu.cn
- 《跟 AI 学 AI 通识》
- 免费兑换码: **gaopin**
24 小时有效!

更多交流

