

AI大模型落地的前景和痛点

兼谈工程师、架构师所面临的机会和挑战

深圳 2023 | 07 | 21

李维 NLP科学家 出门问问VP



“

自我介绍一下

李 维

NLP 老司机、劳碌命，前讯飞AI研究院副院长，出门问问VP。Netbase前首席科学家10年，指挥研发18种语言的社会媒体舆情系统，成为美国NLP产业应用的经典成功案例。Cymfony前研发副总，获第一届问答系统第一名，赢得17个SBIR小企业创新研究项目



Outline



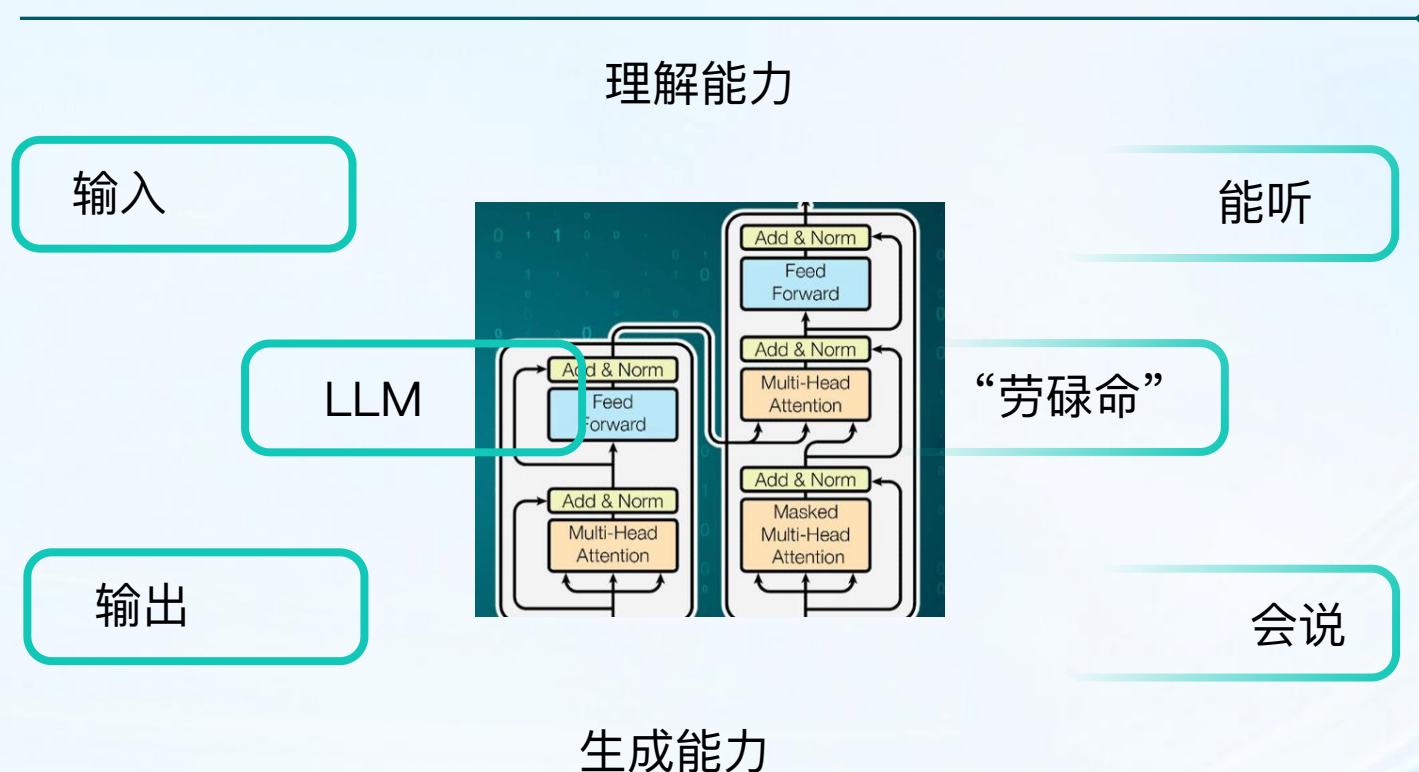
人类语言通天塔的建成



巴别塔建成日：公元2022年11月30日

为什么说建成了人类语言的通天塔？

比 native 还 native



极其简单的原理

自学习、预训练、生成式 AI

(自回归GPT系列)

“大号”的 N-gram 模型

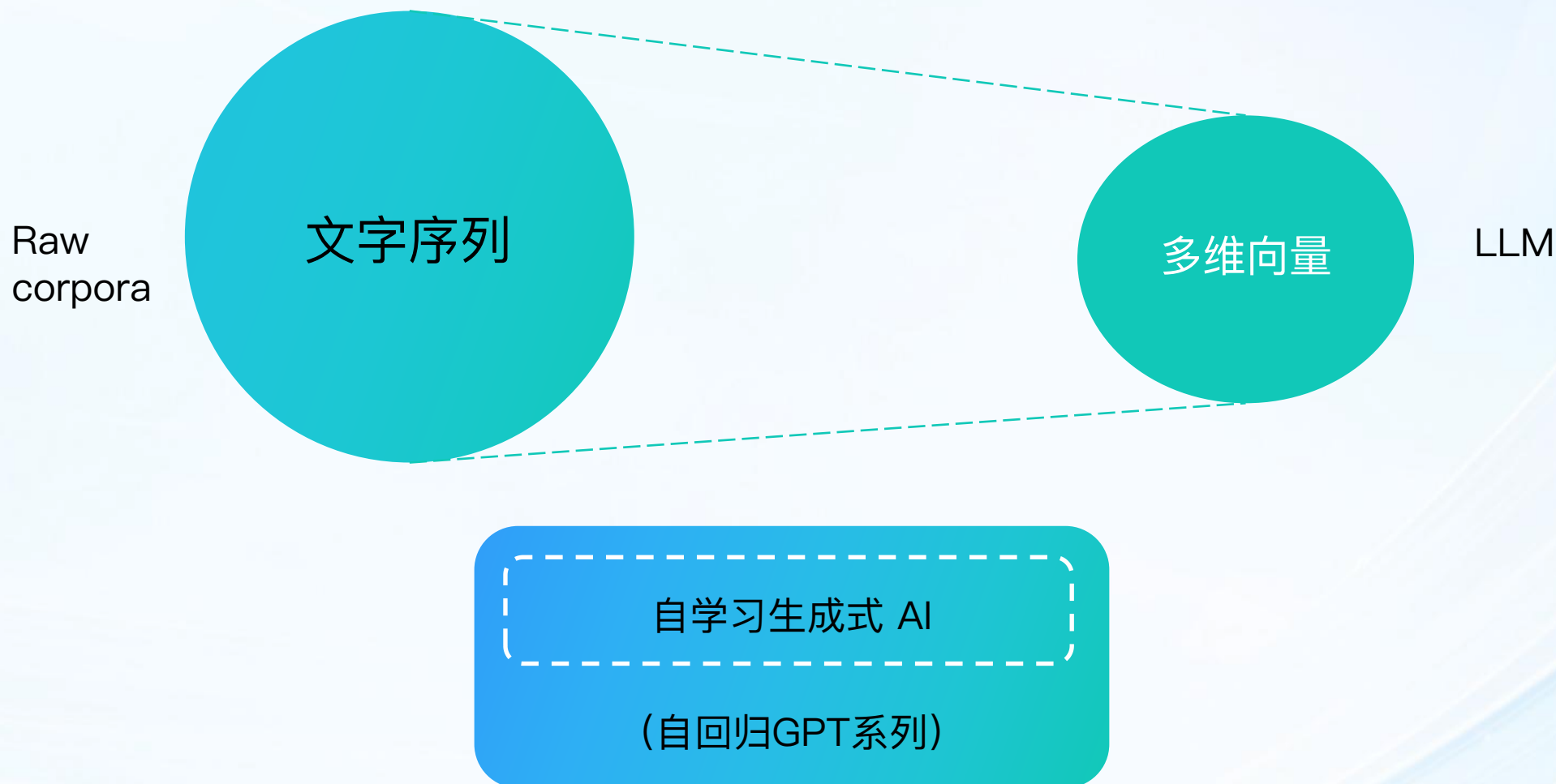
The best thing about AI is its ability to

learn	4.5%
predict	3.5%
make	3.2%
understand	3.1%
do	2.9%

The best thing about AI is its ability to
The best thing about AI is its ability to understand
The best thing about AI is its ability to understand worlds
The best thing about AI is its ability to understand worlds that
The best thing about AI is its ability to understand worlds that are
The best thing about AI is its ability to understand worlds that are
both
The best thing about AI is its ability to understand worlds that are
both exciting
.....

给定上文，预测下一词： next token prediction

大力出奇迹 —— 千（万）亿级字词/千亿级参数



LLM 任务开放

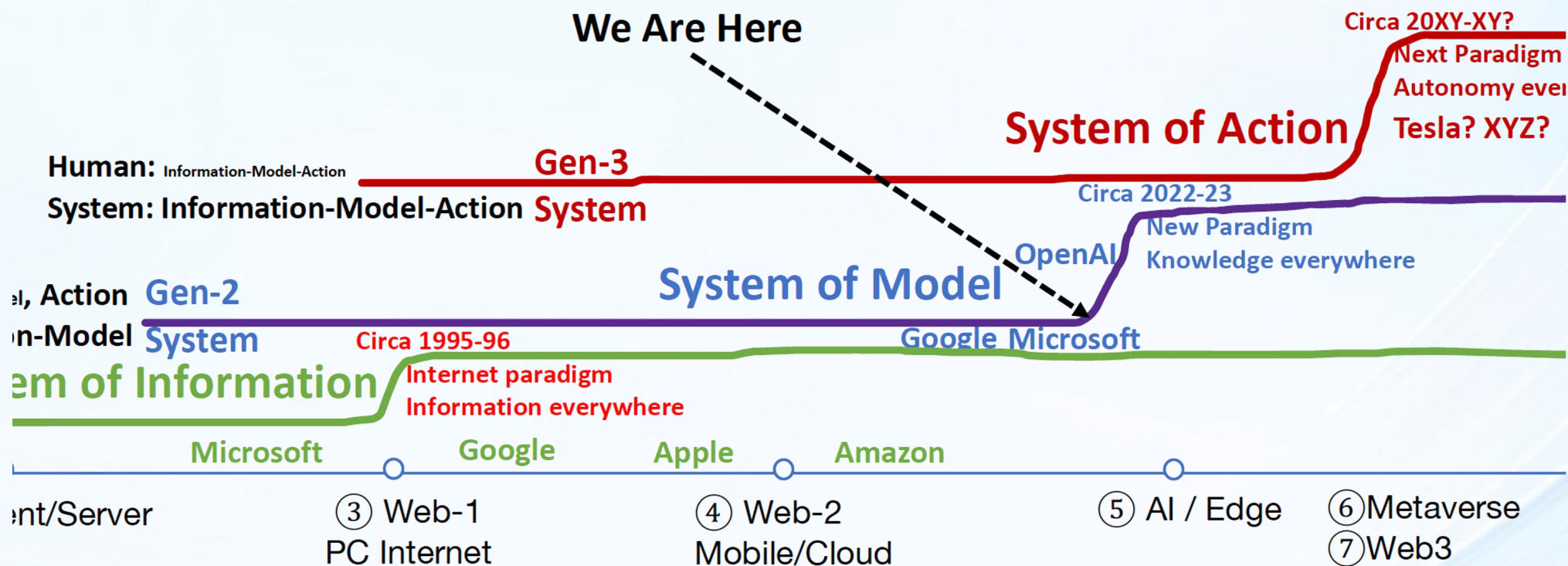
ChatGPT 能做的49件事

问&答	关键字提取	机器学习机器人	产品取名
内容概况	广告设计	文本情绪分析	程序语言转换
程序命令生成	句子简化	生成SQL语句	程序文档生成
Stripe国际API生成	颜色生成	修复代码Bug	代码压缩
结构化生成	段落创作	语言聊天机器人	人称转换
语法纠正	故事创作	清单制作	头脑风暴
生成OpenAI的代码	摘要说明	航空代码抽取	ESRB文本分类
语言翻译	好友聊天	抽取联系信息	点评生成
SQL语句生成	美食制作	文字转表情符号	面试
信息分类	摆烂聊天	程序代码翻译	知识学习
Python代码解释	提纲生成	代码解释	分解步骤
时间复杂度计算	AI聊天	问题解答	高级情绪评分
高级情绪评分	表格填充数据		

Way more
Open-ended

借用陆奇：We Are Here

CF: Web-1 age



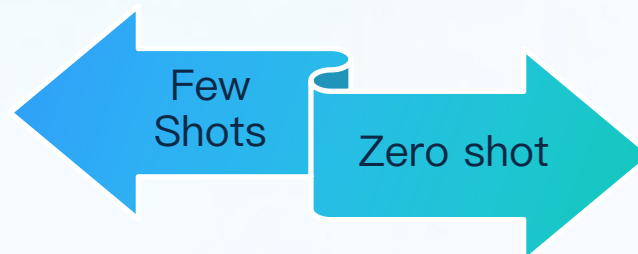
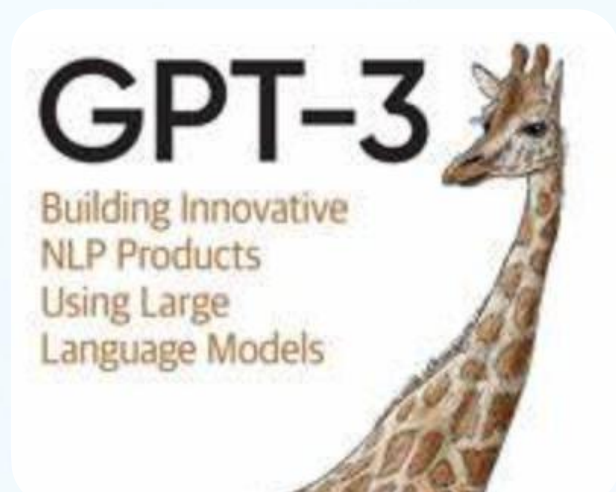
人工智能当代简史

11年前的第一次AI深度神经革命，聚焦感知，靠的是带标大数据有监督训练。

5年多前的二次 AI 深度革命，NLP大模型（LLM）为旗帜，AI从感知到认知，靠的是自监督预训练。

半年前通天塔的三次AI革命，搞定了自然语言交互，以 ChatGPT 为里程碑。

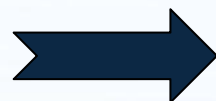
ChatGPT：几乎完美的人机自然语言接口



Zeroshot/零样本：不依赖标注大数据，克服知识瓶颈

机器迁就人，而不是人迁就机器

海量样本标注



少样本标注



零样本标注

ChatGPT 能力测试



Zeroshot 开放类任务理论上没有测试集，无法定量数据质量 (QA)

客户调查

用户粘性、增长性

利用已有测试：
NLP测试集，高考，各资格考试

机器迁就人的结果：人皆码师

码农贵族、产品经理/CXO、数据分析员、用户群体：

底线拉平，LLM 听得懂所有人的指示

为“万众创业”创造了条件

NLP“核爆炸”及其影响

新突破导向新生态

ChatGPT 建成语言通天塔

- 语言大模型搞定人类语言
- 机器迁就人，而不是人迁就机器
- 实体交互方式的革命：不仅仅是人机
- 语言模型蕴含了知识模型，虽然并不完备

GPT 洗礼后的新生态建设

- 标志感知智能走向认知智能：语言+知识
- LLM CoPilot 落地，AIGC 前景广阔
- 用户接口：交互克服人工智障
- LLM Agent 研究值得关注

大模型时代的挑战

模型痛点、领域挑战

模型痛点

- 幻觉问题
- 稳定性问题
- 知识欠缺问题
- 多模态问题

领域落地的挑战

- 谁调用谁：插件与外挂
- 领域数据库、业务逻辑、私有数据
- 自主AI研究：Agent
- 与现有领域 Legacy 系统的关系

大模型时代的焦虑

架构师/工程师焦虑：you are not alone!

NLP 老司机的焦虑

- 一夜醒来，专业没了
- 有说：半年前的专业经验是负资产
- AI 系统的涌现，资源的爆发
- 百花齐放，AI 日新月异

普罗的焦虑

- 工作被取代？
- 艺术被颠覆？
- 知识更新跑不赢知识增长

“幻觉”： 是 LLM 内在 feature 不是 bug

幻觉与想象力是同义词

预训练的 feature, not a bug

- 没有“幻觉”，如何讲“故事”（cf 《人类简史》）
- 写小说，不是报道新闻：记者成不了小说家
- 长尾细节=噪音，抗噪=遗忘细节
- 模型不是数据库：源于数据，高于数据
- 张冠李戴是顺畅生成的必然
- 知道自己不知道？ 一切都是概率
- 是遗忘，不是“说谎”： 没学会人类的语气词



对于部分落地场景的挑战

- “一正胡八”： 说谎不“脸红”，真假莫辨，极度误导
- 有利于艺术创造任务，却是领域落地的命门
- 加大模型可以减少幻觉，不能根治幻觉
- 强化学习可能医治“一本正经”，不能根治“胡说八道”
- 领域场景解决方案： 外挂数据库
- 人来把关和核实： 可以先假设细节都是不真实的

“不稳定”： 是 LLM 内在 feature 不是 bug

不稳定与创造性是同义词

预训练的 feature, not a bug

- 不稳定，才有创造性、多样性
- 生成式LLM的本性：概率模型
- 不稳定=随机采样，太稳定=枯燥重复
- 模型不是数据库：是条件预测，不是检索
- 不稳定带来惊喜或惊恐

对于部分落地场景的挑战

- 容易做 demo，不容易独立落地
- 有利于艺术创造任务，却也是领域落地的命门
- 加大模型可以缓解不稳定，不能根治不稳定
- 多次生成：majority vote
- 人来把关和挑选：副驾驶模式，亦可使多次生成

知识欠缺： 是 序列学习LLM之本性

LLM 不是数据库

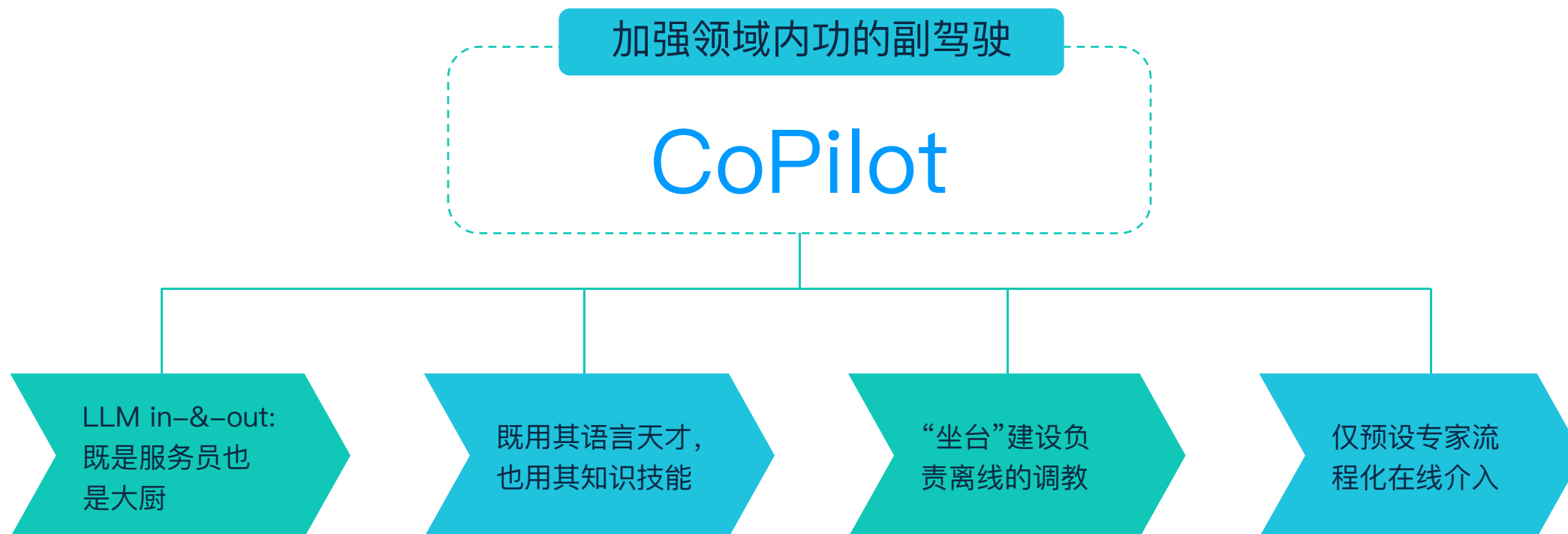
序列数据预训练的边界

- 知识渊博，尤其是百科知识和多数常识
- 领域知识欠缺
- 模型不是数据库：是条件预测，不是检索

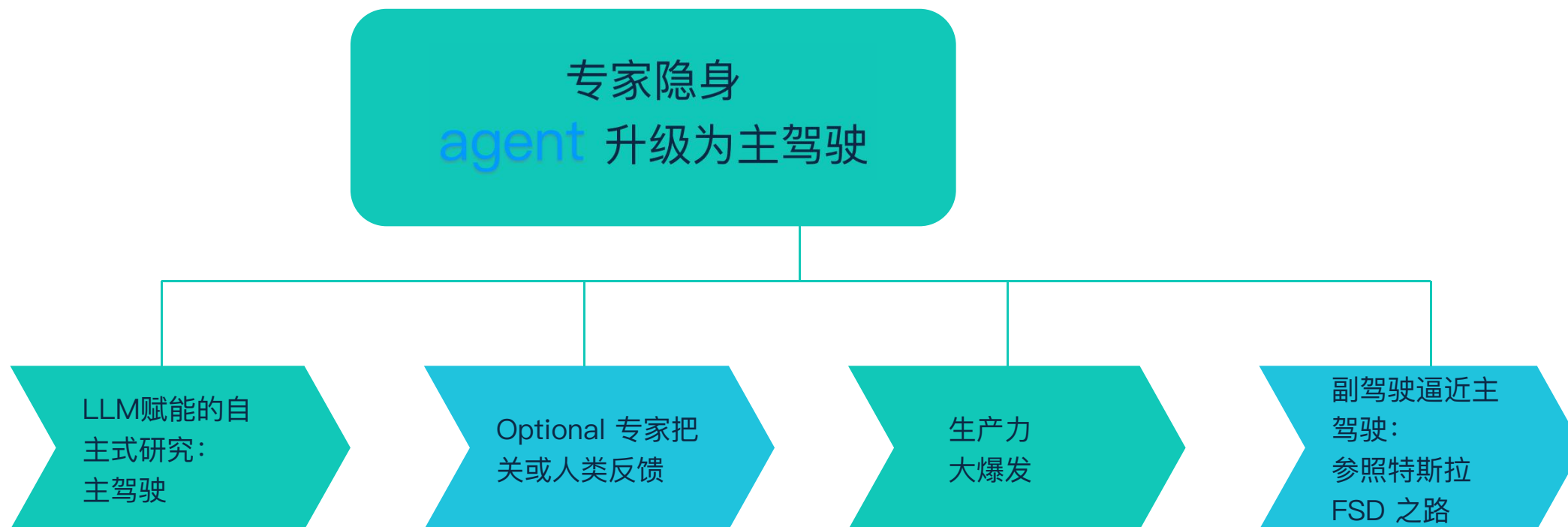
垂直领域的知识短板

- 容易做 demo，不容易独立落地
- 加大模型可以增加知识，不能穷尽知识
- 业务逻辑
- 提示工程：in-context learning, short memory
- 传统领域数据库、知识库、术语词典、知识图谱
 - 向量数据库很火
 - 解决背景信息不足
 - long memory

专家坐台的打造



LLM既是后台又坐前台



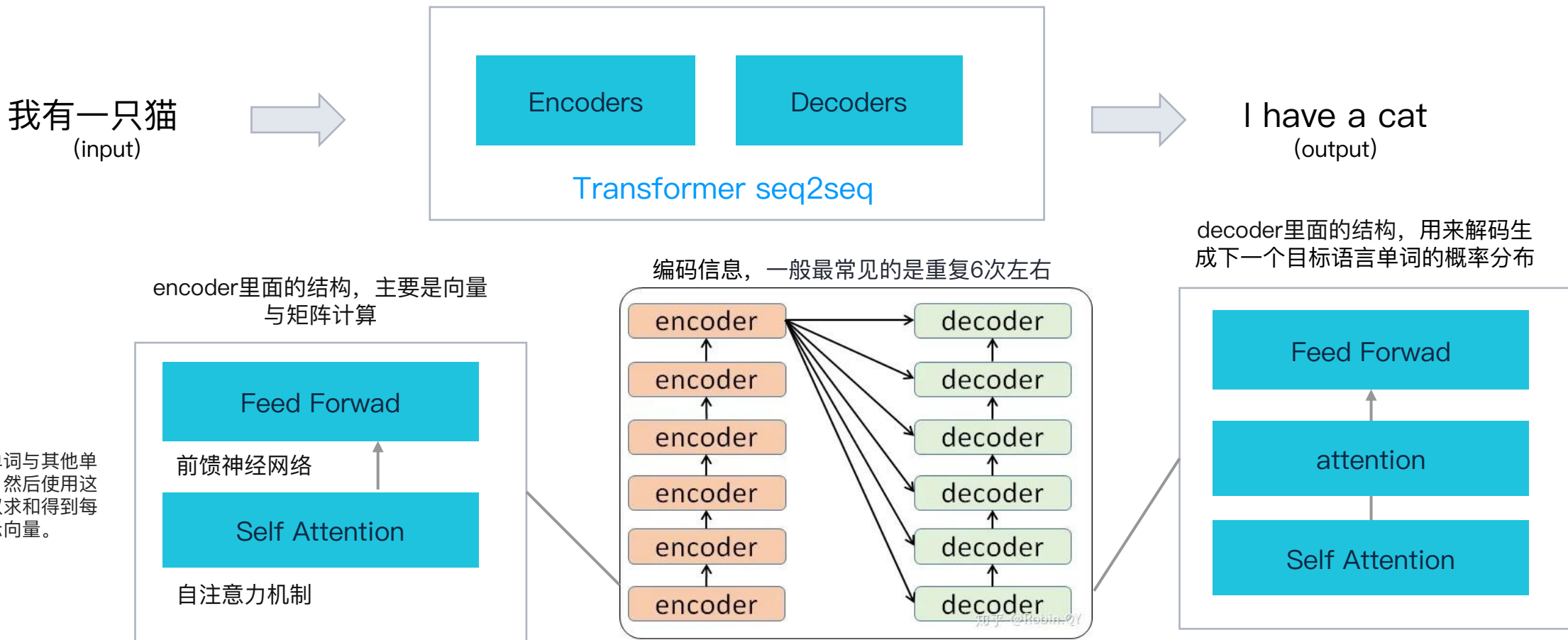
2



大模型预训练及其微调机制

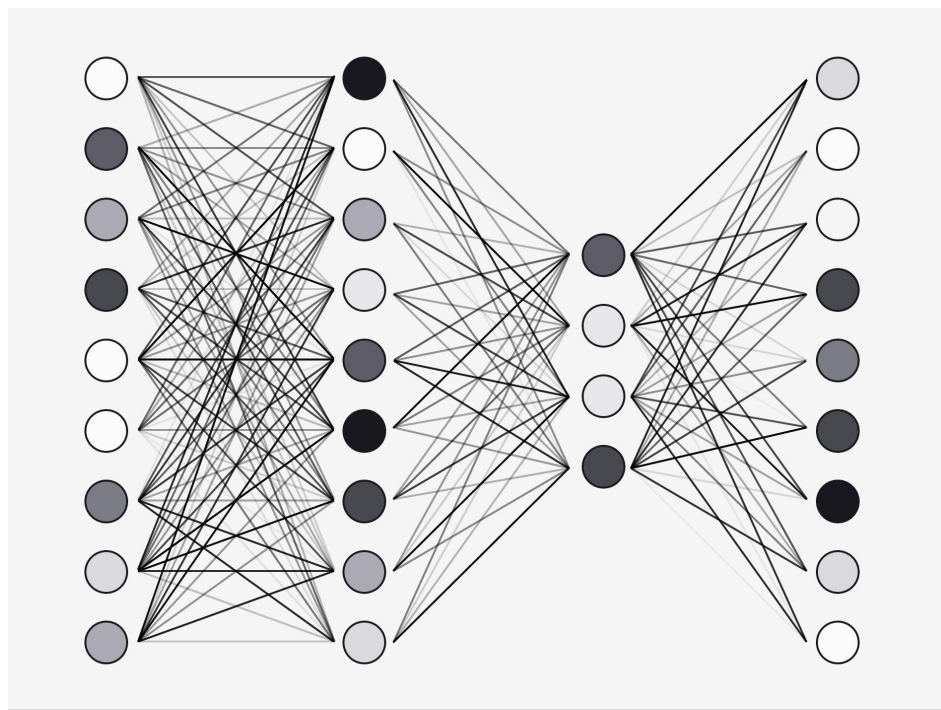
Transformer架构

完全基于注意力机制，完全不用递归和卷积
特别擅长对序列数据中的长距离依赖进行建模，非常适合自然语言处理任务



参数（权重）

大语言模型，本质上是一个包含了大量参数的神经网络。这个神经网络可以看作是一个复杂的函数，它接收输入（如一段文本），然后生成输出（如下一个词）。



神经网络的每一层都可以看作是一个函数，它接收上一层的输出（或者输入数据，如果它是第一层），然后生成自己的输出。

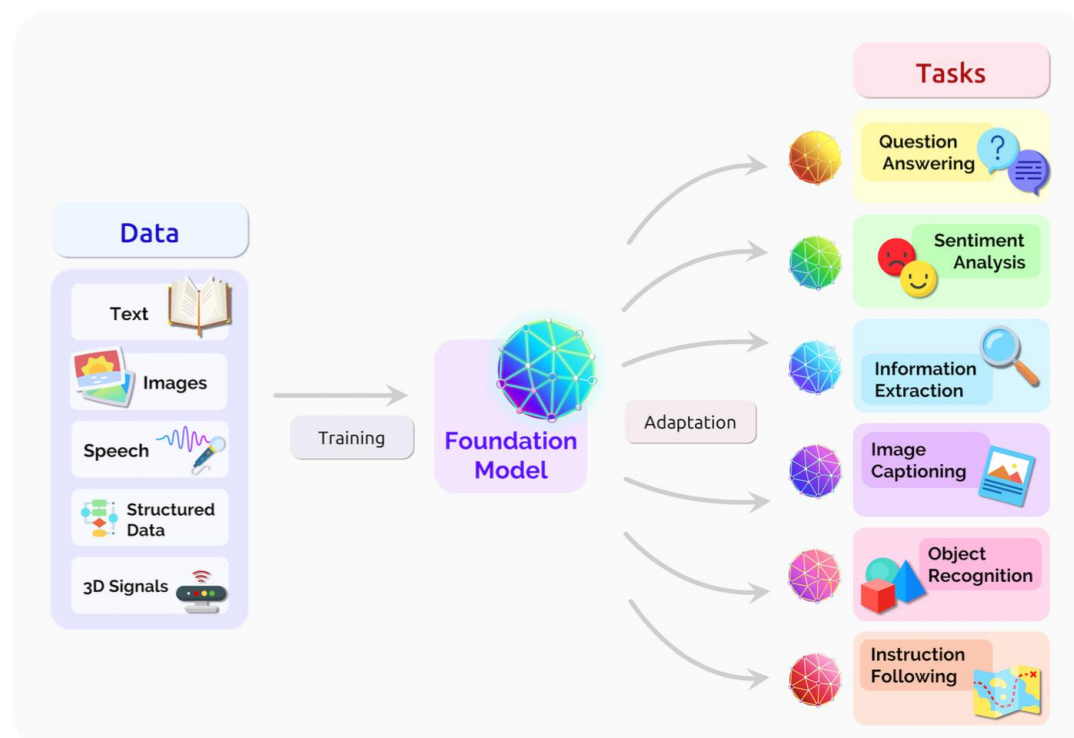
每个层都包含许多的神经元。每个神经元都有一个权重，这就是大语言模型的参数。这些权重在训练过程中被调整，以使模型能够更好地预测输出。

大模型之预训练

AI大模型就是Foundation Model（基础模型），指通过在大规模宽泛的数据上进行训练后能适应一系列下游任务的模型。

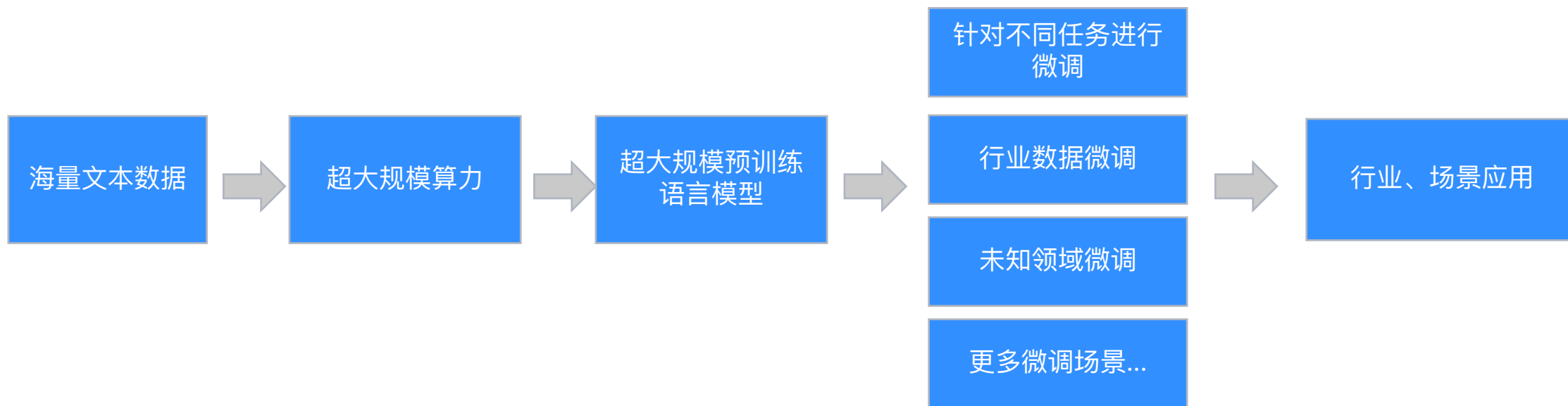
GPT-3训练数据来自于互联网上的大量文本，包括维基百科、新闻文章、网页、论坛、社交媒体等。

模型先进行预训练，即在未标记的文本上进行多次迭代的自监督学习，以学习语言的结构和规律。



“大规模预训练 + 微调”范式提升模型通用性

“预训练+微调”技术，可用一套技术解决不同语言、不同场景、不同 NLP 任务，有效地提升了开发效率。



微调方式：任务对齐

大模型底座更通用，顶层更垂直，fine-tune是在不改动预训练模型的基础上，在模型「顶层」使模型能够更贴合实际使用场景。

来源 Stanford Alpaca: A Strong,
Replicable Instruction-Following Model



来源：《follow instructions with human
feedback》



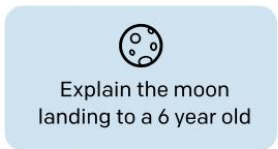
大模型底座

基于人类反馈强化的训练：偏好对齐

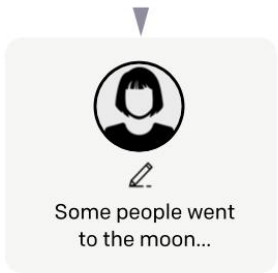
通过人工标注对模型输出结果打分建立奖励模型，然后通过奖励模型继续循环迭代，OpenAI获得了更真实、更无害，并且更好地遵循用户意图的语言模型InstructGPT。

步骤1：搜集说明数据，训练监督策略
有监督微调（SFT）

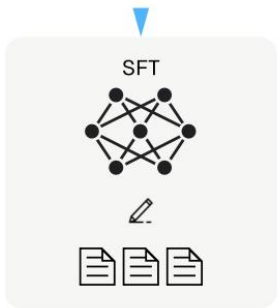
搜集说明数据(基于prompt训练方式的数据集)，训练监督策略



Labeler(标注人员)揭示期望的输出行为

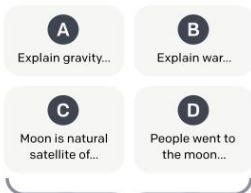
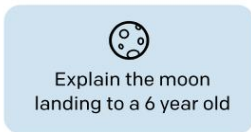


数据用来联合监督学习，对GPT-3进行微调

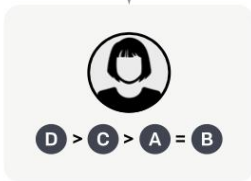


步骤2：搜集比较数据，训练一个奖励模型
奖励模型（RM）训练

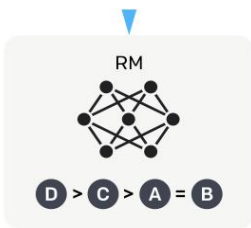
对这个prompt训练的数据集和若干模型的结果进行抽样



Labeler(标注人员)从最优到最差将输出结果进行排序



这个数据用来训练反馈模型



步骤3：搜集说明数据，使用增强学习优化模型
通过PPO根据奖励模型进行强化学习。

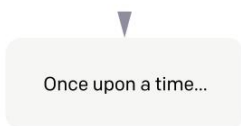
新的prompt从数据集中抽样



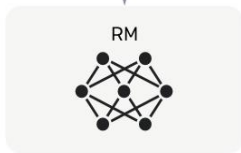
借助模型生成输出



反馈模型为输出计算一个反馈结果



反馈结果用来优化策略



3



“序列猴子”大模型 及其AIGC 产品矩阵

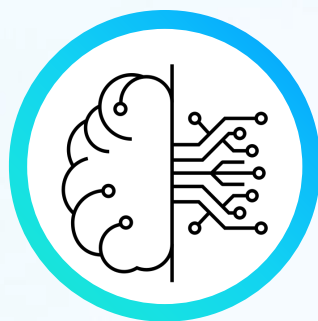
介绍一下出门问问的大模型与AIGC的模式

三位一体的独特商业模式 飞轮效应凸显



CoPilot将会无处不在

每个人/企业都应该有专属的“ChatGPT”



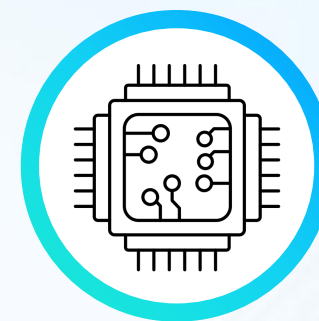
高智商的大脑



动听的声音



好看的外表



智能硬件载体

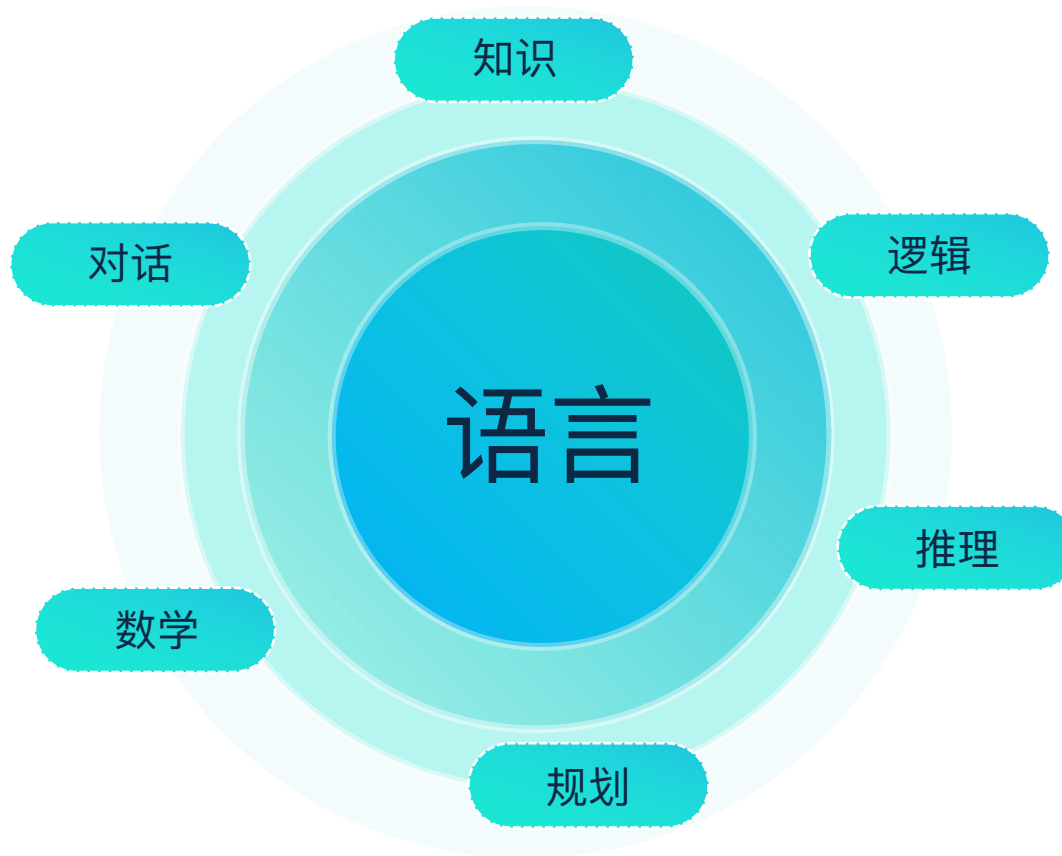
出门问问大模型：序列猴子-Sequence Monkey

无限猴子定理

只要给予无限时间，一只猴子几乎必然能够打出任何文字，比如莎士比亚的《哈姆雷特》

序列猴子“定理”

如果喂给机器猴子海量的文本序列，通过大规模算力去训练它，就能练就有智能的“序列猴子”



大模型的能力体系



序列猴子

序列猴子：面向创作者的应用

出门问问研发了多款面向创作者的AIGC产品及应用，为万千创作者构建一站式CoPilot产品矩阵，打通内容创作全流程。



奇妙文



言之画



魔音工坊



奇妙元

超600万

注册会员

超40万

付费会员

奇妙元—您的AI数字分身

像做PPT一样制作数字人视频和直播

多元化高清数字人 3种形态

海量资产

- ◆ 200+数字形象
- ◆ 形象克隆
- ◆ 3D捏脸
- ◆ IP激活

丰富语音

- ◆ 1000+音色
- ◆ 2000+声音风格
- ◆ 40国语言

逼真效果

- ◆ 立体面部
- ◆ 4K高清
- ◆ 超长序列



本视频由出门问问AIGC应用
“奇妙元”通过一张照片生成
声音由“魔音工坊”生成。



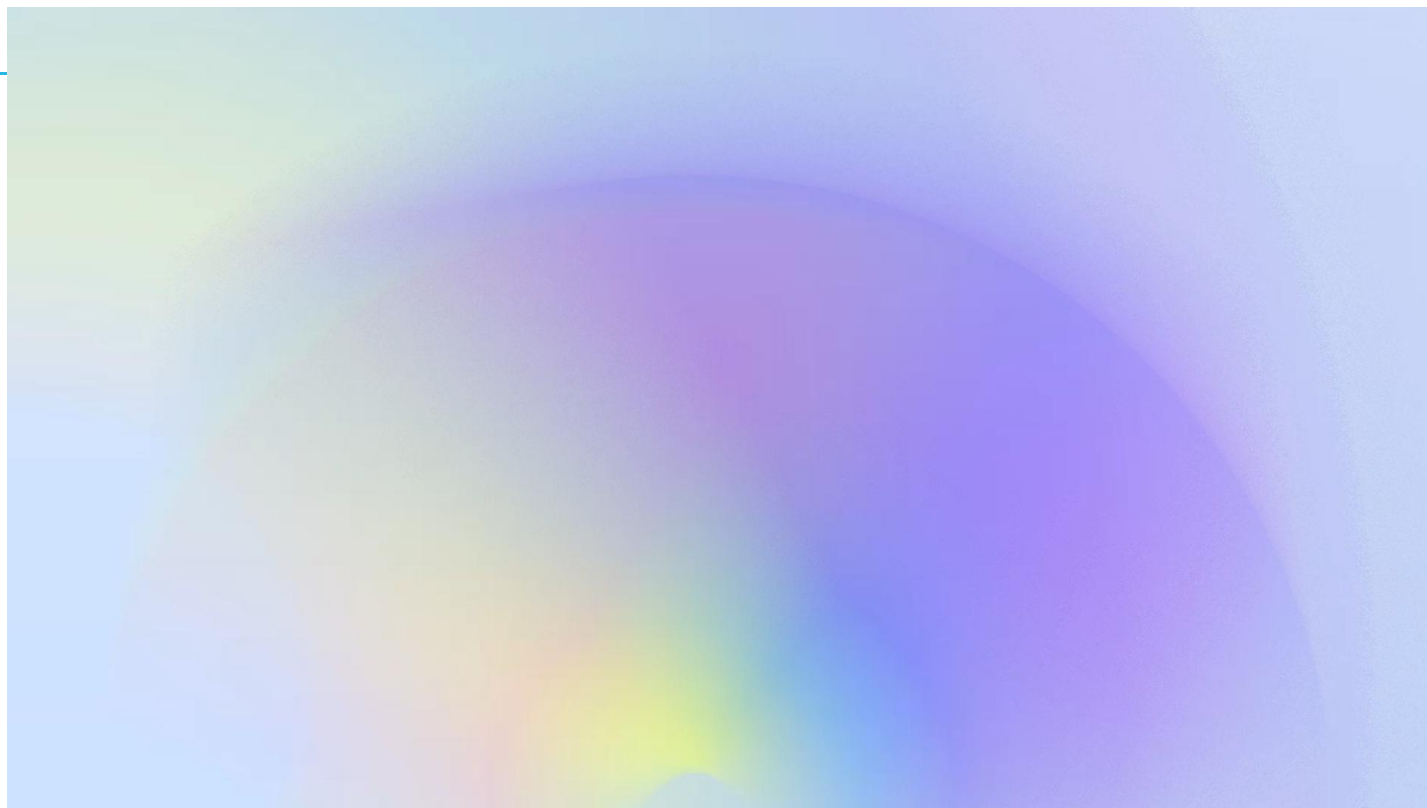
25年前的 yours truly

本视频由出门问问AIGC应用
“奇妙元”通过一张照片生成
声音由“魔音工坊”生成。



如今的“我”： 劳碌命

本视频由出门问问AIGC应用
“奇妙元”通过形象克隆生成
声音由“魔音工坊”生成。



AIGC产品内测探索通道



奇妙文

write.mobvoi.com



言之画

paint.mobvoi.com



魔音工坊

moyin.com



奇妙元

weta365.com

魔法小问(App)

大模型与语音助手结合，实现更强、更智能的功能。

魔法小问：“Siri”+“ChatGPT”



大模型大脑

- 语音助手
- 个人信息记忆
- 文档分析
- 图片生成



数字人形象

- 好看的外表
- 好听的声音



企业专属大模型定制

序列猴子大模型，开放生态赋能各行各业



企业专属大模型开发提供的服务

出门问问为企业客户提供专属的大模型定制化开发服务，确保服务全流程打通与高质量交付。



企业专有数据

01 模型定制

根据客户的需求和场景，为客户量身定制一个适合其业务和数据的大模型。包含API接口定制、数据定制、模型架构与参数定制

02 数据处理

帮助客户在使用大模型之前，对其数据进行清洗、预处理、转换和规范化等，以确保数据的质量和一致性，并使其符合大模型的输入要求

03 模型训练

使用企业的数据对模型进行训练，并进行模型评估。
帮助客户在其自身数据上训练大模型，以满足企业特定的业务需求

04 私有化部署

针对特定行业或企业内部业务场景定制的大模型，在企业内部进行部署和应用，以满足客户个性化的应用需求，同时保证数据隐私和安全

05 参数调优

帮助客户在训练好的大模型上进行参数调试，以获得更好的性能和准确性。

商业模式

出门问问公有云

01

通用API

◆ 通用数据

◆ 通用模型

通用API接口，快速、稳定

02

专属API+
专属CoPilot

◆ 企业专有数据

◆ 通用模型

构建向量数据库

03

专属模型
API

◆ 企业专有数据

◆ 企业私有模型

私有化数据，训练专属模型

04

初级私有化
部署

◆ 企业专有数据

◆ 通用模型

通用模型+CoPilot部署

05

高级私有化
部署

◆ 企业专有数据

◆ 企业私有模型

私有化数据，训练专属模型

企业私有云

前沿动态/商务合作



公众号：飞哥说AI



大模型招聘：LLM@mobvoi.com