

《人工智能通识》（科技素养）

第10讲 人工智能前沿

主讲：丛润民

章节知识点概览



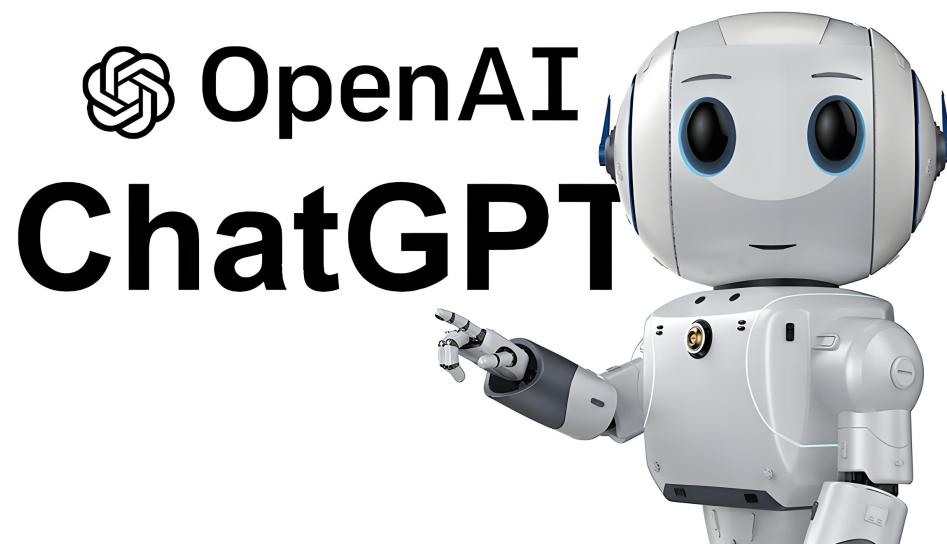
- 知识点1：以规模铸就优势——大模型的兴起与演进
- 知识点2：AI新纪元的全能王者——多模态大模型
- 知识点3：从基础大模型到垂域大模型
- 知识点4：“AI”上创作
- 知识点5：全新的科学发现范式——AI for Science
- 知识点6：玩转大模型

以规模铸就优势——大模型的兴起与演进



01 大模型的前世今生

02 大模型的训练与优化



CH 如何看待世界互联网大会这个平台?

世界互联网大会是一个重要的国际平台，专门讨论互联网技术及其对社会、经济和文化的影响。这个平台汇聚了来自世界各地的互联网专家、技术人员、政府官员和商界领袖，共同探讨互联网发展的未来趋势和挑战。

世界互联网大会的重要性在于，它促进了全球互联网技术和应用的发展，促进了国际间的技术交流和合作。它还为政府、企业和其他组织提供了一个展示最新互联网技术和创新成果的平台，

CH 你能帮我写个类似于“羊了个羊”的小程序代码嘛

当然可以！这是一份用 Python 实现的“羊了个羊”小程序代码：

```
python
import random

def sheep_game():
    # 设置羊数量
    sheep_count = random.randint(10, 100)

    # 设置循环次数
    loop_count = 0

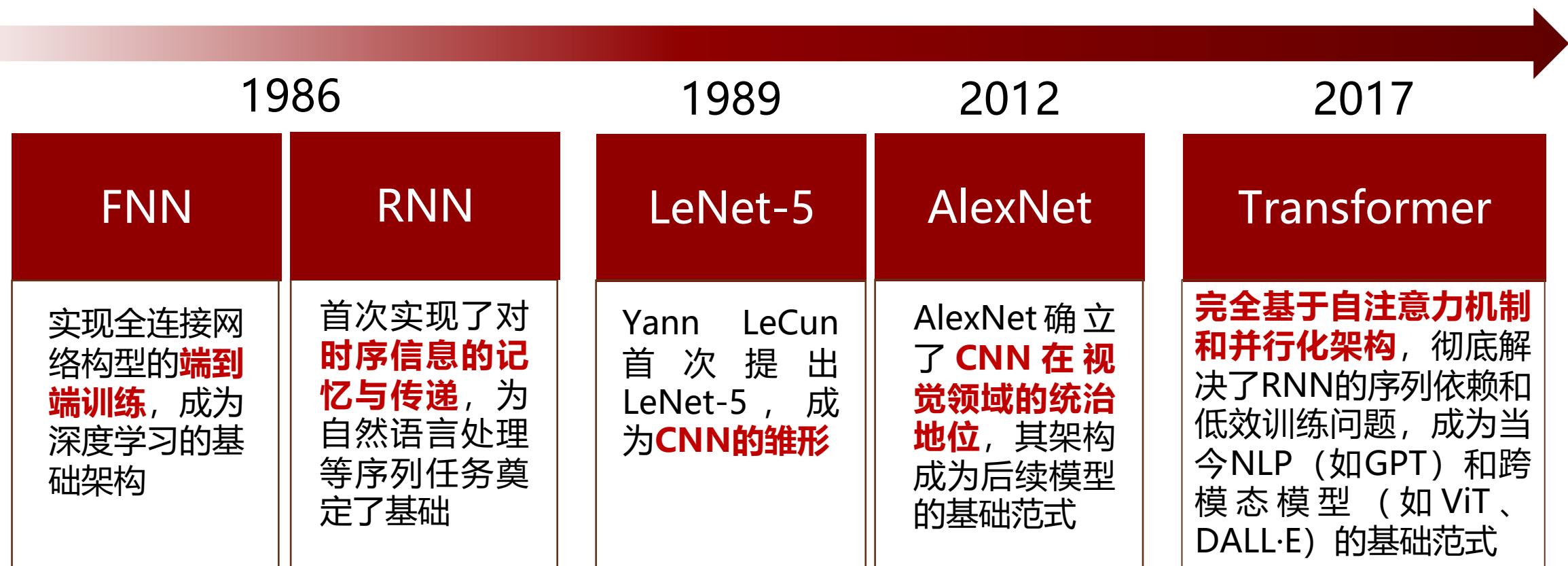
    while sheep_count > 0:
        print("羊数量: ", sheep_count)
        sheep_count -= 1
        loop_count += 1

    print("经过了", loop_count, "轮羊了个羊，所有羊都被羊完了。")
```

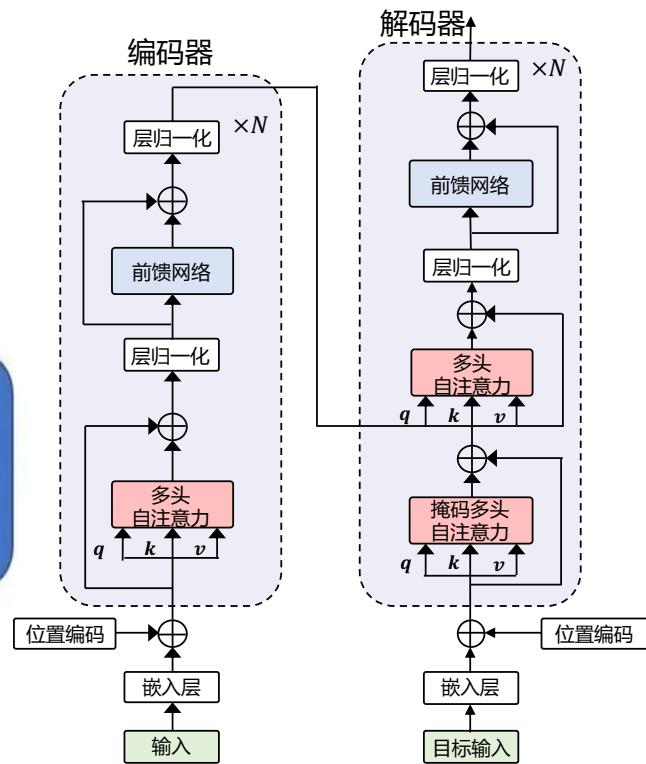
6分钟看懂
什么是
知识增强大模型



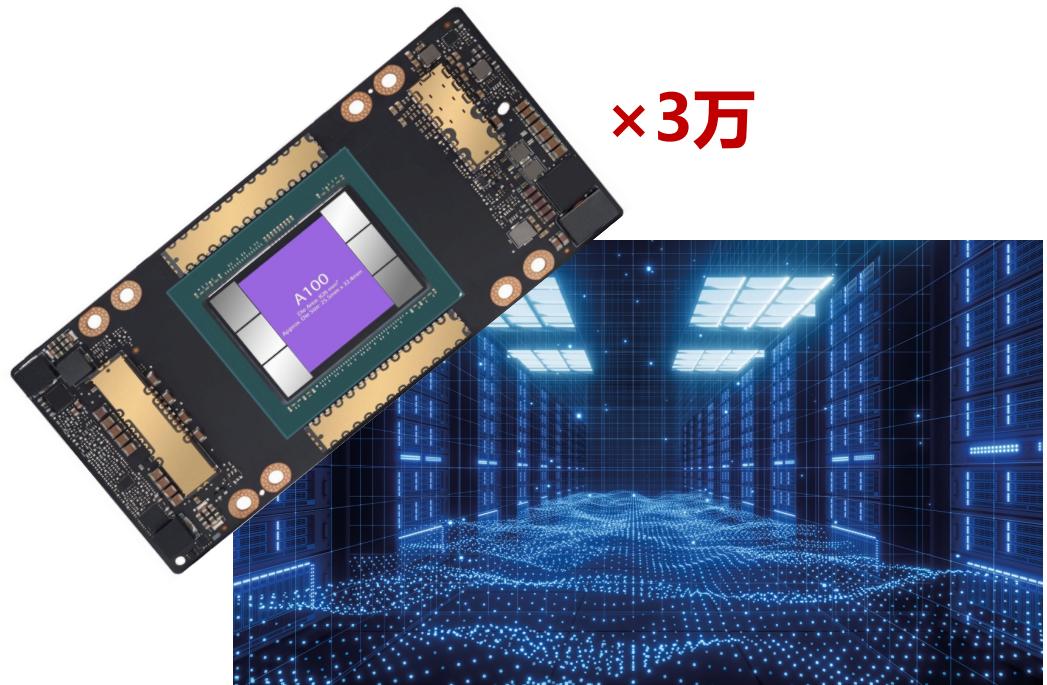
■ 谷歌提出了Transformer架构，利用自注意力机制构建长程序列关系的同时，实现了对模型的并行化训练，这为研发大模型提供了可以并行优化的基础模型结构。



- 谷歌提出了**Transformer架构**，利用自注意力机制构建长程序列关系的同时，实现了对模型的并行化训练，这为研发大模型提供了可以并行优化的基础模型结构。



ChatGPT使用Transformer为核心组件



训练一个ChatGPT的模型
需要上万张A100！！！



GPT家族发展史

2018

2019

2020

2022

2023

GPT

GPT-2

GPT-3

InstructGPT

ChatGPT
(GPT-3.5)

GPT-4

有一定泛化能力，用于监督任务无关的NLP任务中

生成能力显著增强：
聊天、续写、
编故事

绝大部分任务：
代码生成、模仿叙事、
生成剧本等

提示学习与基于人类反馈的学习

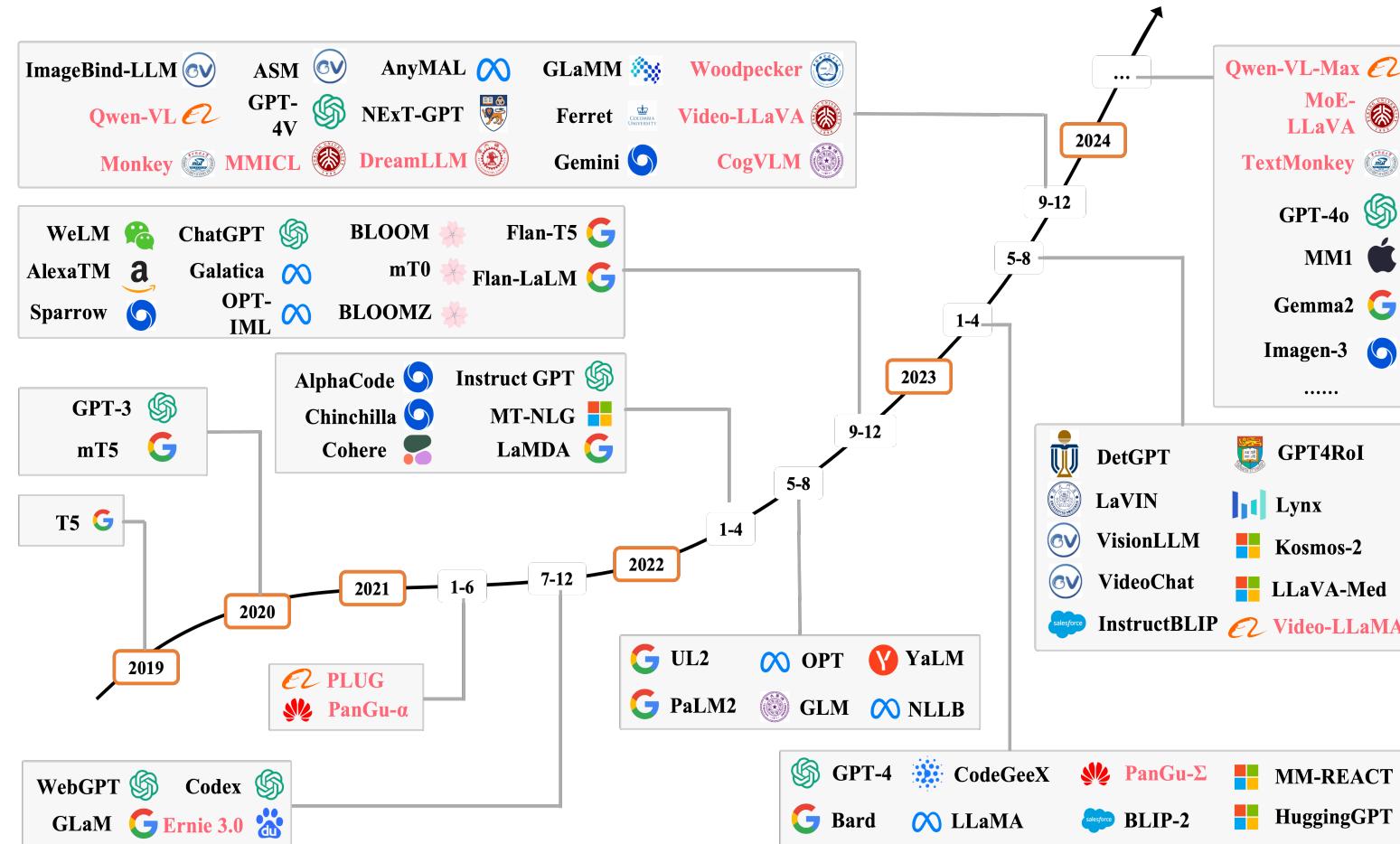
增强版
InstructGPT

多模态等
更强能力

彻底“出圈”

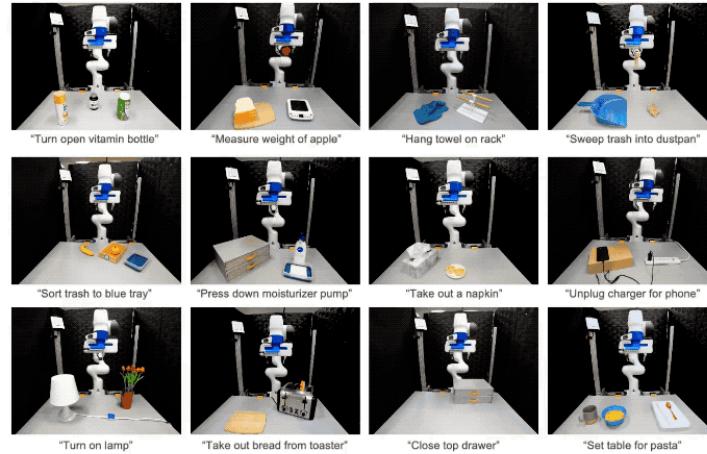


■ 大模型，顾名思义，是指那些参数规模巨大、计算量复杂的深度学习模型。



规模持续扩大提升性能
多模态融合增强应用能力
行业定制化满足垂直需求

■ 大模型的优势与挑战皆根植于其“**大**”这一本质特征之中。



➤ **零样本迁移能力：**模型在没有见过特定任务的训练数据的情况下，仅根据已有的知识和对其他相关任务的学习经验，就能对该新任务进行处理和做出合理预测的能力。而且大模型通常可以在多个任务之间迁移和共享知识，这意味着在特定任务上训练的大模型可以迅速适应其他相关任务。



➤ **涌现能力：**随着模型规模的增大，模型表现出一些原本在较小规模时没有明显出现的新功能或特性。这些能力往往不是通过明确的编程或者训练目标获得的，而是在模型的规模达到一定程度后自发涌现的，称为“涌现能力”。

■ 大模型的研发主要分为两个阶段，即预训练阶段和微调与对齐阶段。

➤ 预训练是大模型研发的关键阶段，其成效直接关乎模型后续能力的广度与深度。预训练阶段首先要准备的便是大规模、高质量的训练数据。

例句：张明正在\$#&学习大模型的相关内容。张明正在学习大模型的相###关内容。



■ 大模型的研发主要分为两个阶段，即预训练阶段和微调与对齐阶段。

- 在模型架构方面，目前国内外主流的大模型主要基于Transformer框架包括以下3种：
- 编码器-解码器 (Encoder-Decoder) —— "翻译官模式"

输入句子：[天气][真][好]

↓ ↓ ↓ (编码器全部可见，蓝色)

编码记忆： 

↑ ↑ ↑ (解码器可随时查看，红色)

输出生成：[The] [weather] [is] [nice]

→ 生成时能参考完整输入

- 因果解码器 (Causal Decoder) —— "写日记模式"

生成顺序：[今天][天气][真][好]

↓ ↓ ↓ ↓

可见范围：[今]—[今天]—[今天天气]—[今天天气真]
(绿色可见，灰色被遮挡)

■ 大模型的研发主要分为两个阶段，即预训练阶段和微调与对齐阶段。

➤ 在模型架构方面，目前国内外主流的大模型主要基于Transformer框架包括以下3种：

- 编码器-解码器 (Encoder-Decoder) —— "翻译官模式"
- 因果解码器 (Causal Decoder) —— "写日记模式"
- 前缀解码器 (Prefix Decoder) —— "命题作文模式"

前缀提示：[写科幻故事：] [火星上...]

 (蓝色，始终可见)

生成内容：[人类][发现了][金字塔]



可见范围：[火]—[火星上]—[火星上人类]—[火星上人类发现了]

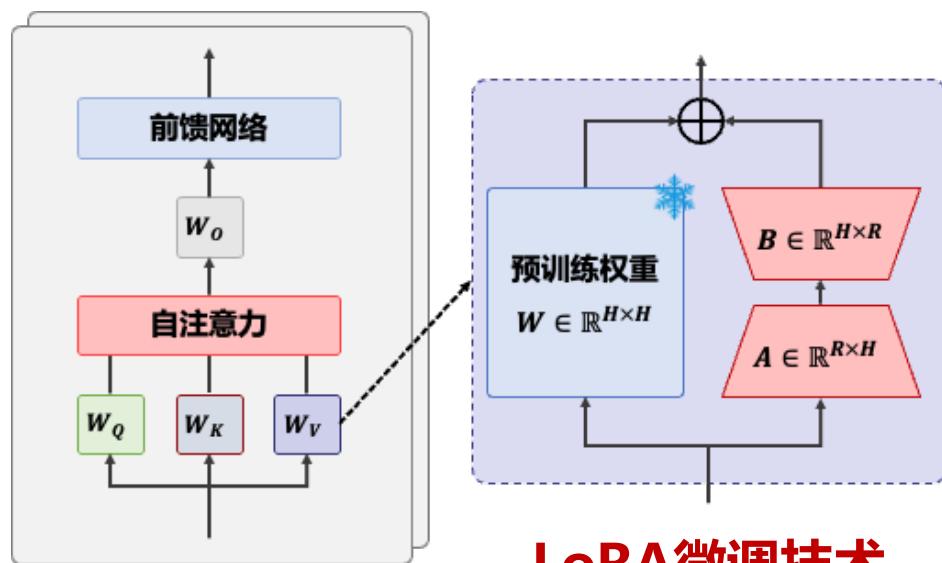
(绿色部分仍受因果限制)

■ 在进行大规模预训练时，往往需要设计合适的**自监督预训练**任务，使得模型能够从海量无标注数据中学习到广泛的语义知识与世界知识。

	语言建模	去噪自编码器	混合去噪器
目标	学习词元序列的概率分布	学习从破坏的文本中恢复原始文本	学习从多种破坏模式中恢复原始文本
输入	完整的文本序列	被破坏的文本	被多种方式破坏的文本
输出	无	删除、替换、重排词元等	多种噪声类型（如删除、替换、重排等）
噪声强度	无	固定或可调	可调
模型结构	编码器	编码器+解码器	编码器+解码器
训练目标	最大化下一个词元的预测概率	最小化重构文本与原始文本的差异	最小化重构文本与原始文本的差异
损失函数	负对数似然	均方误差或交叉熵损失	均方误差或交叉熵损失
应用场景	序列生成、翻译、问答等	文本补全、错误纠正、文本恢复等	多种文本处理任务，增强鲁棒性和泛化能力
鲁棒性	较低	较高	更高
泛化性	中等	较强	更强
优点	学习词元间依赖关系	提高模型对抗噪声的能力	提高模型对抗多种噪声的能力
缺点	可能忽视上下文信息	只能处理特定类型的噪声	训练过程可能很复杂

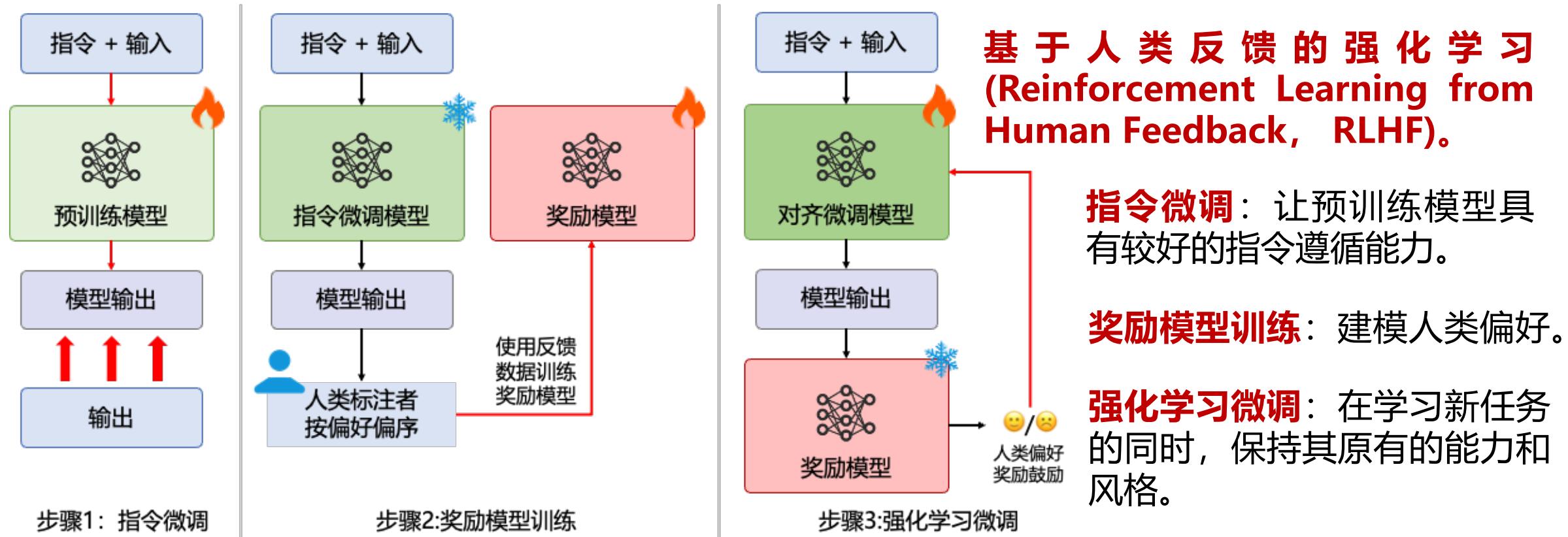
- 在特定的应用场景下，需要对大模型进一步优化，使其能够有效地应用预训练阶段获取的知识，并理解、适应人类意愿，在不同任务下表现出优秀的准确性和适用性，从而更好地服务于实际应用需求。

指令微调：收集或构建指令化的实例，然后通过有监督的方式对大语言模型的参数进行微调。



LoRA微调技术：在不修改原有模型权重的情况下，通过引入低秩矩阵来调整模型的某些层，从而实现快速适应新任务的目的。

- 在特定的应用场景下，需要对大模型进一步优化，使其能够有效地应用预训练阶段获取的知识，并理解、适应人类意愿，在不同任务下表现出优秀的准确性和适用性，从而更好地服务于实际应用需求。



- 目前大模型的性能依然在持续不断地刷新，并不断有新的大模型推出，如火爆出圈的“国货之光”Deepseek大模型，堪称大模型的性价比典范。



遥遥领先的国产大模型之光



DeepSeek



KO



OpenAI

以超高的性价比一夜之间火爆全球！

知识点2： AI新纪元的全能王者——多模态大模型



01 多模态大模型的概念与演进

02 CLIP——多模态大模型的先驱

多模态大模型

多模态大模型 简直就是AI界的"六边形战士"！



即梦AI

字节跳动推出的一站式AI创作平台



扣子PPT

免费一键生成精美PPT



AiPPT

AI快速生成高质量PPT



美图设计室

AI图像创作和设计平台



绘蛙

AI电商营销工具，免费生成商品图



办公小浣熊

超强AI数据分析助手



HitPaw

专注于AI视频、图像和音频处理工具



Opera Neon

Opera推出的全新AI Agent浏览器



找我呀

本地AI知识助手，文件搜索和智能问...



讯飞绘文

免费AI写作工具，5分钟生成一篇原创...



笔灵AI写作

面向专业写作领域的AI写作工具



新华妙笔

新华社推出的体制内办公学习平台



Paperpal

英文论文写作助手



笔目鱼

专业英文论文写作者



稿易AI论文

AI论文写作助手，免费生成2000字大纲



沁言学术

AI科研写作平台，一站式文献管理...

今天呢我们聊一个AI圈的关键词



光速写作

AI写作、PPT生成工具，单篇章长150...



华文算杆

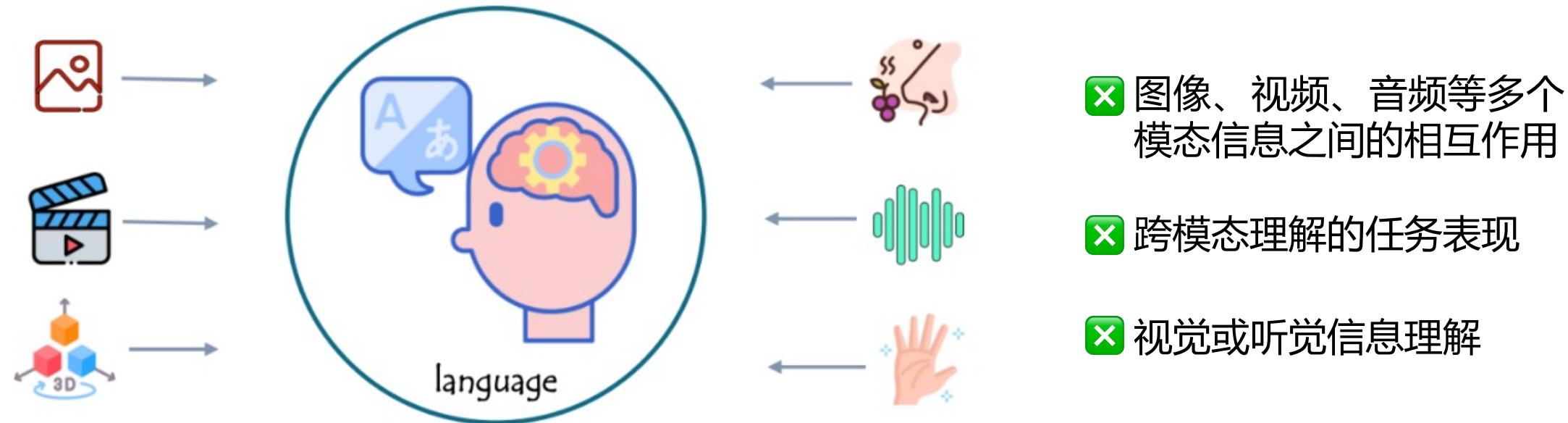


秘塔写作猫



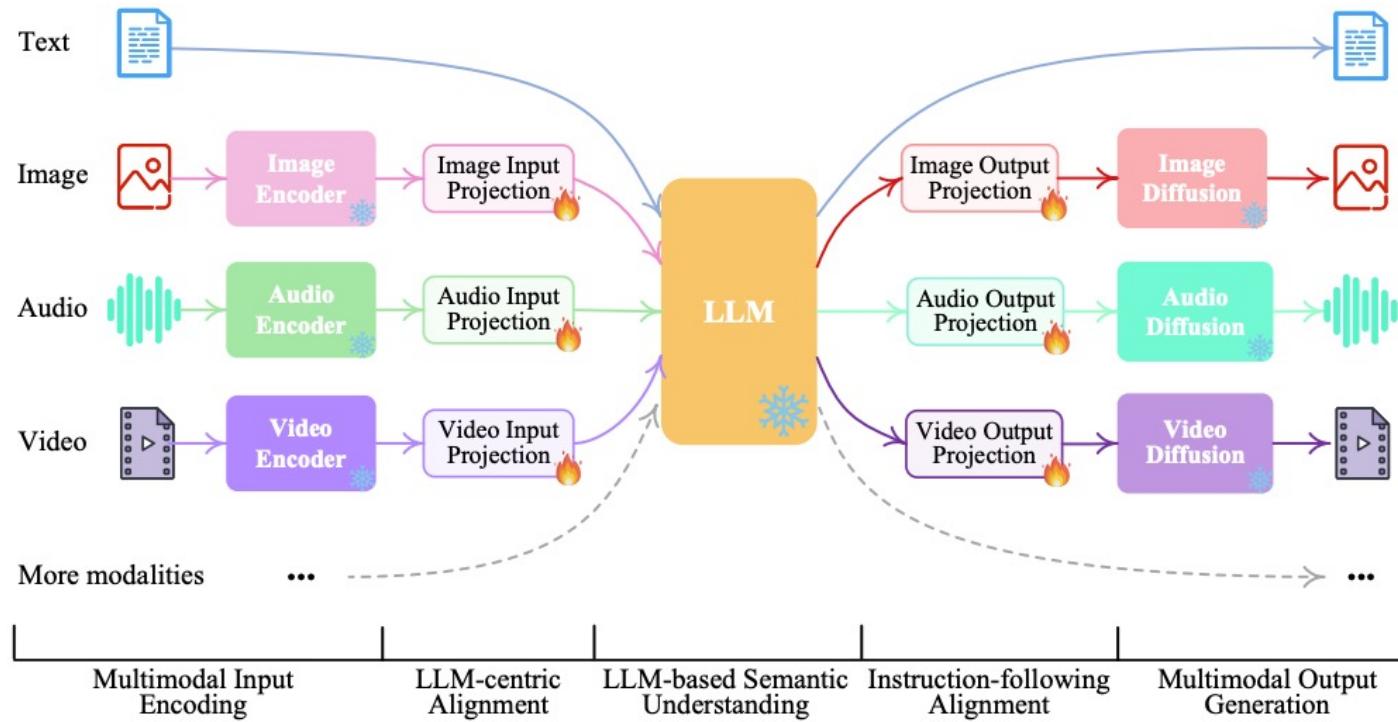
讯飞写作

■ 大语言模型，与其他模态的模型连接起来，实现协同对推理，催生了一个新的研究领域——**多模态大语言模型（Multi-Modal Large Language Model, MLLM）**



多模态大语言模型通过对多种模态数据的联合学习和推理，模型能够捕捉到不同模态之间的内在关联和互补信息，从而更全面地理解和描述现实世界中的事物和现象。

- 多模态大模型的关键能力在于整合并理解不同的数据格式，应用于需要理解和整合不同类型数据信息的任务。**LLMs处于核心位置，多模态在LLMs基础上进行扩展**，即找到一个方法将不同模态的数据映射到LLMs可以接收的语义空间。



多模态编码器：负责将不同模态的输入数据编码为模型可理解的表示。

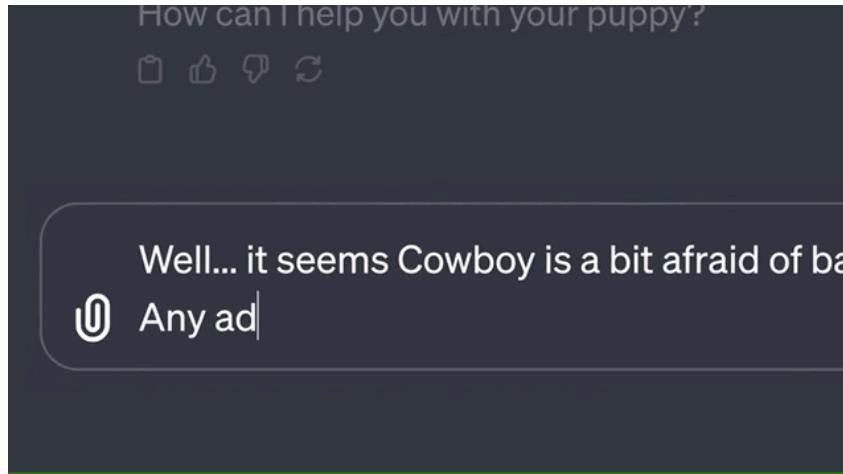
输入映射：将不同模态的输入数据映射到共享的语义空间。

LLMs：大型语言模型，用于处理文本数据。

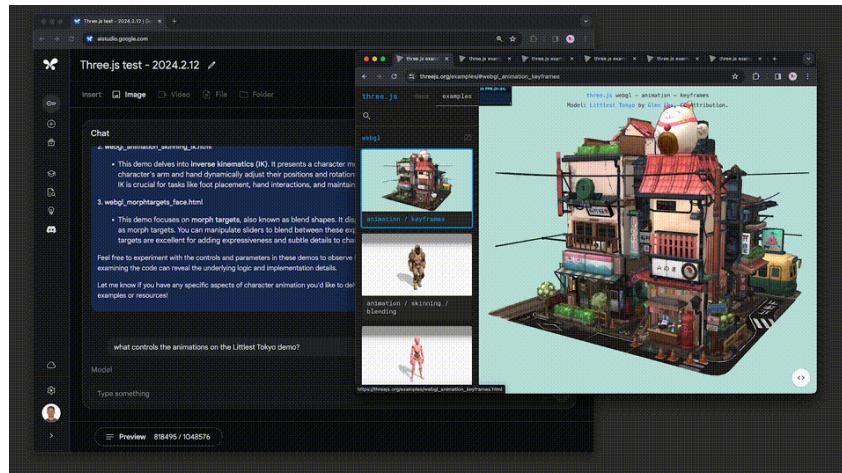
输出映射：将模型生成的输出映射回原始模态的空间。

多模态生成器：根据输入数据生成对应的输出数据。

多模态大模型的概念与演进



GPT-4理解图片信息充当宠物助手



Gemini进行3D场景重建

多模态内容理解和文本生成



特定模态内容生成

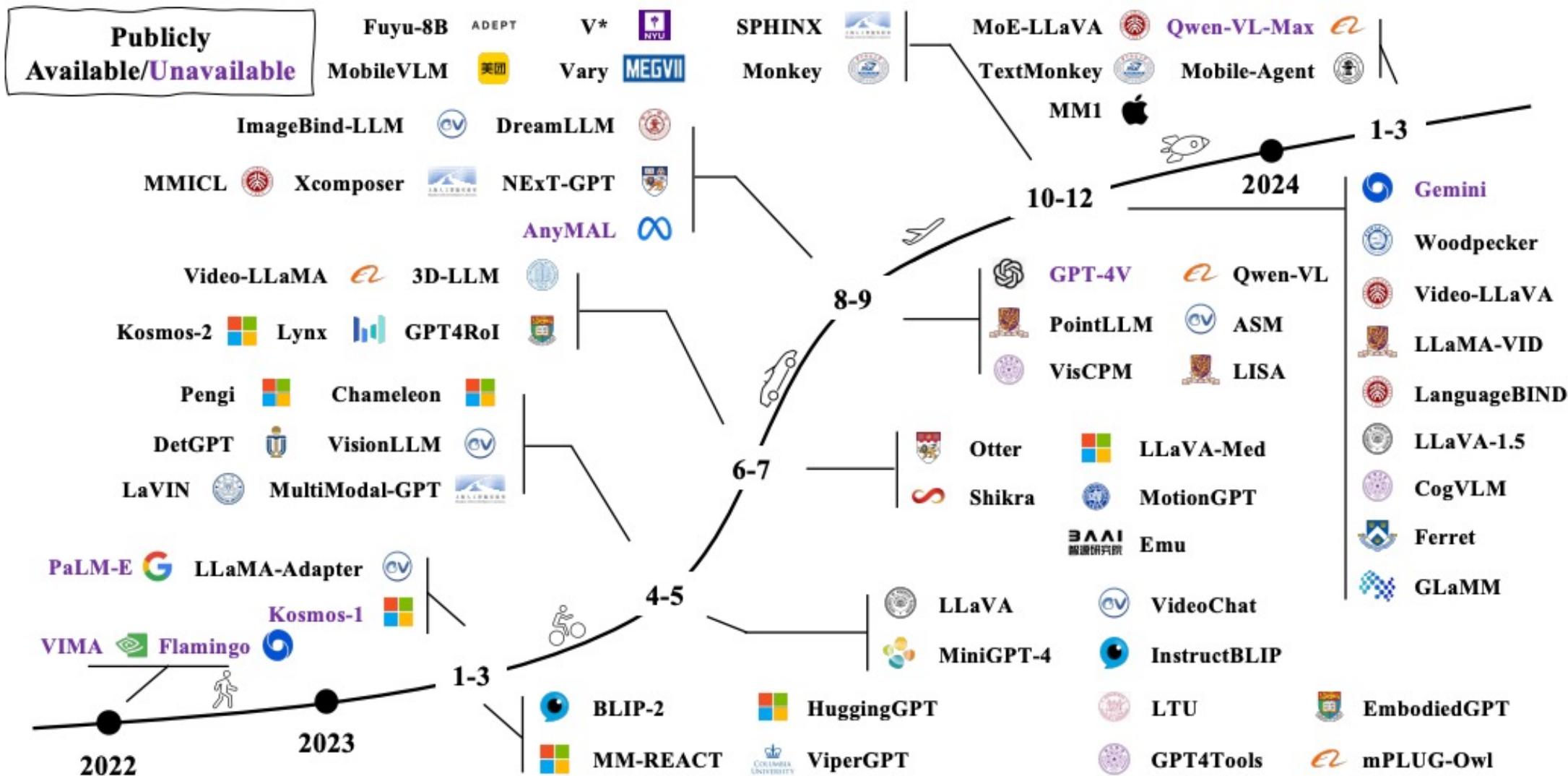


特定模态内容生成

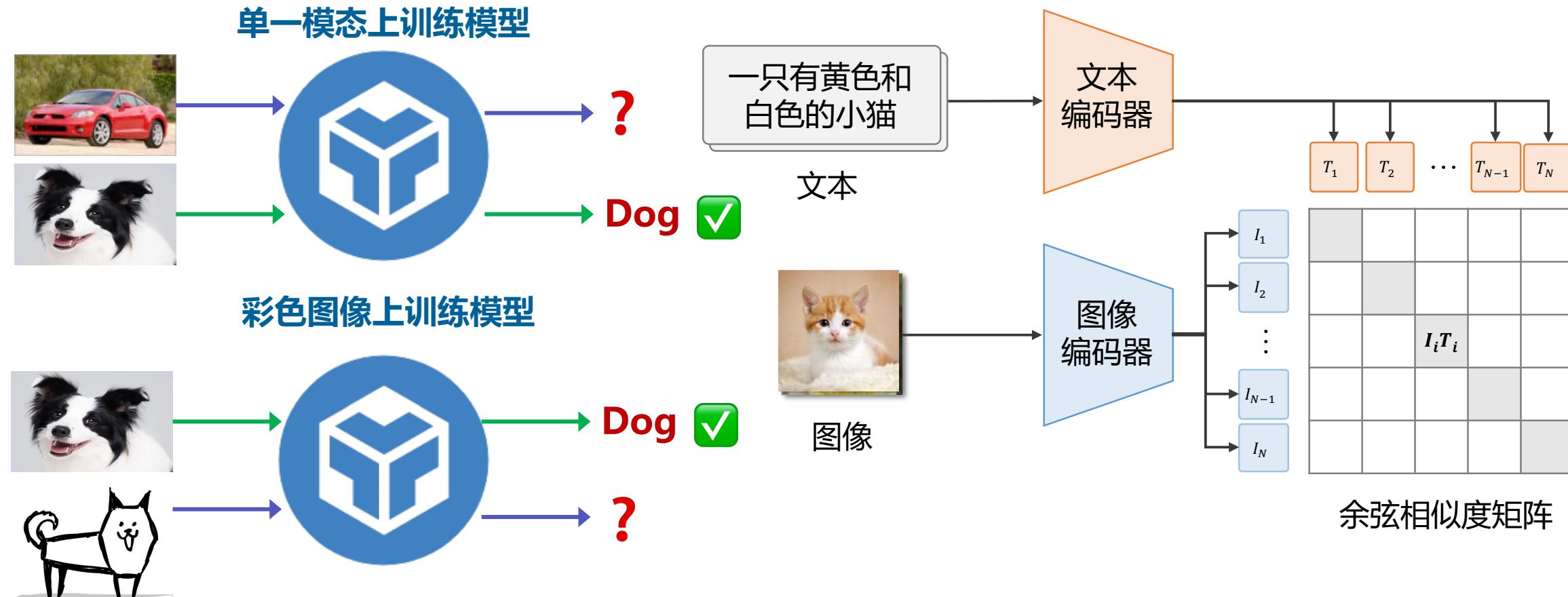


多模态大模型的概念与演进

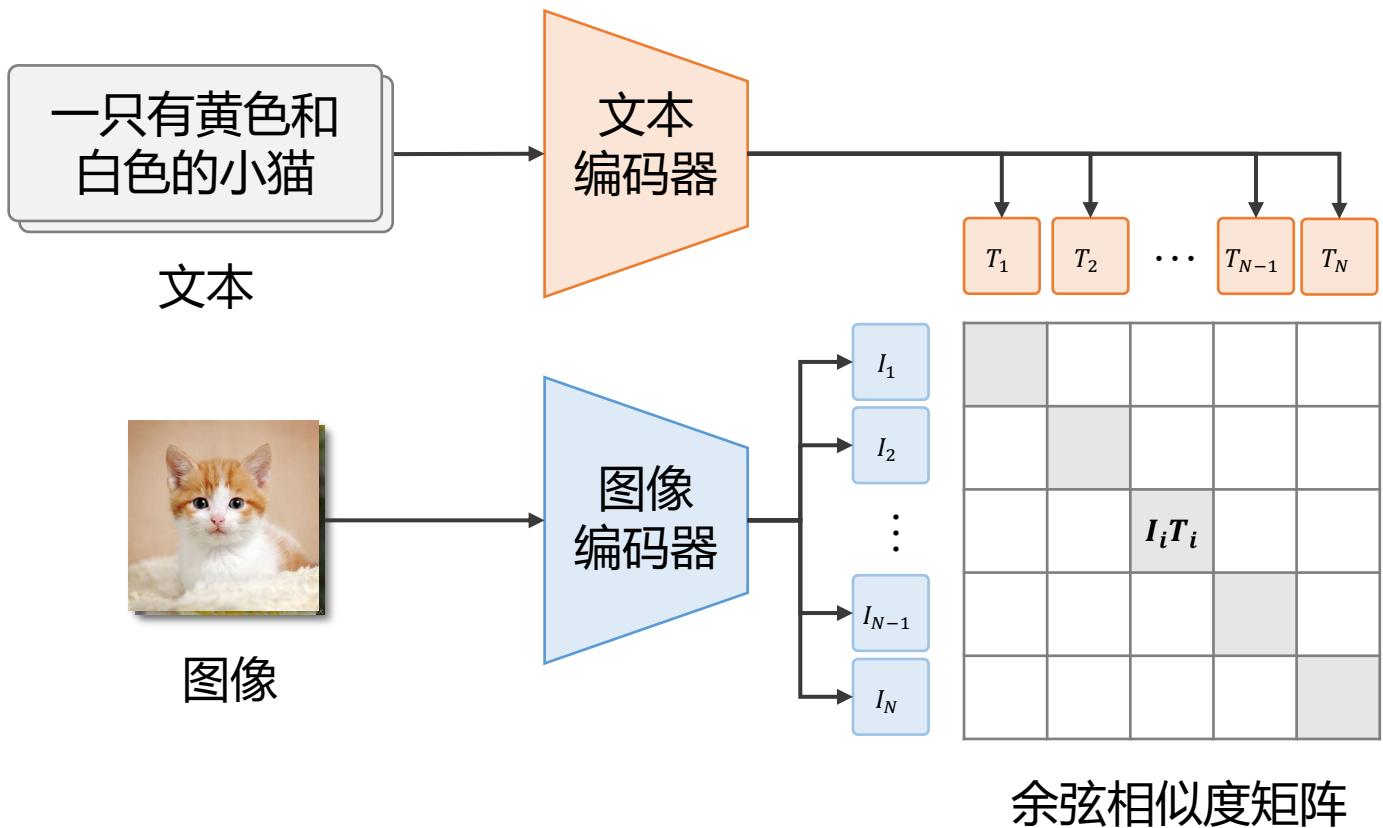
■ 多模态大模型历经融合发展，如今其综合处理能力不断提升，在各行业广泛应用。



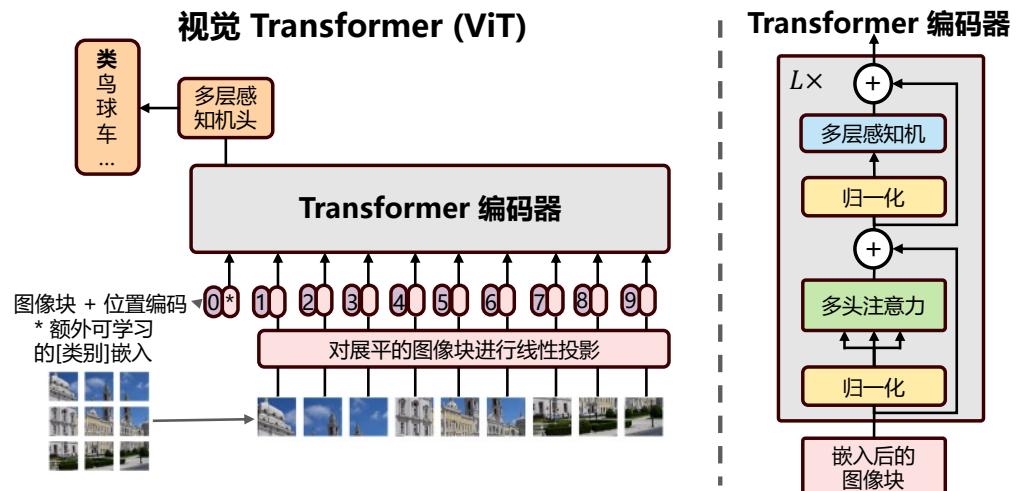
■ CLIP (Contrastive Language-Image Pre-training) 探索了如何利用大规模未标注数据进行有效的多模态学习，在图像分类、检索和生成等任务上表现出了强大的迁移能力，为多模态大模型的发展提供了重要的技术借鉴和思路启发。



CLIP——多模态大模型的先驱

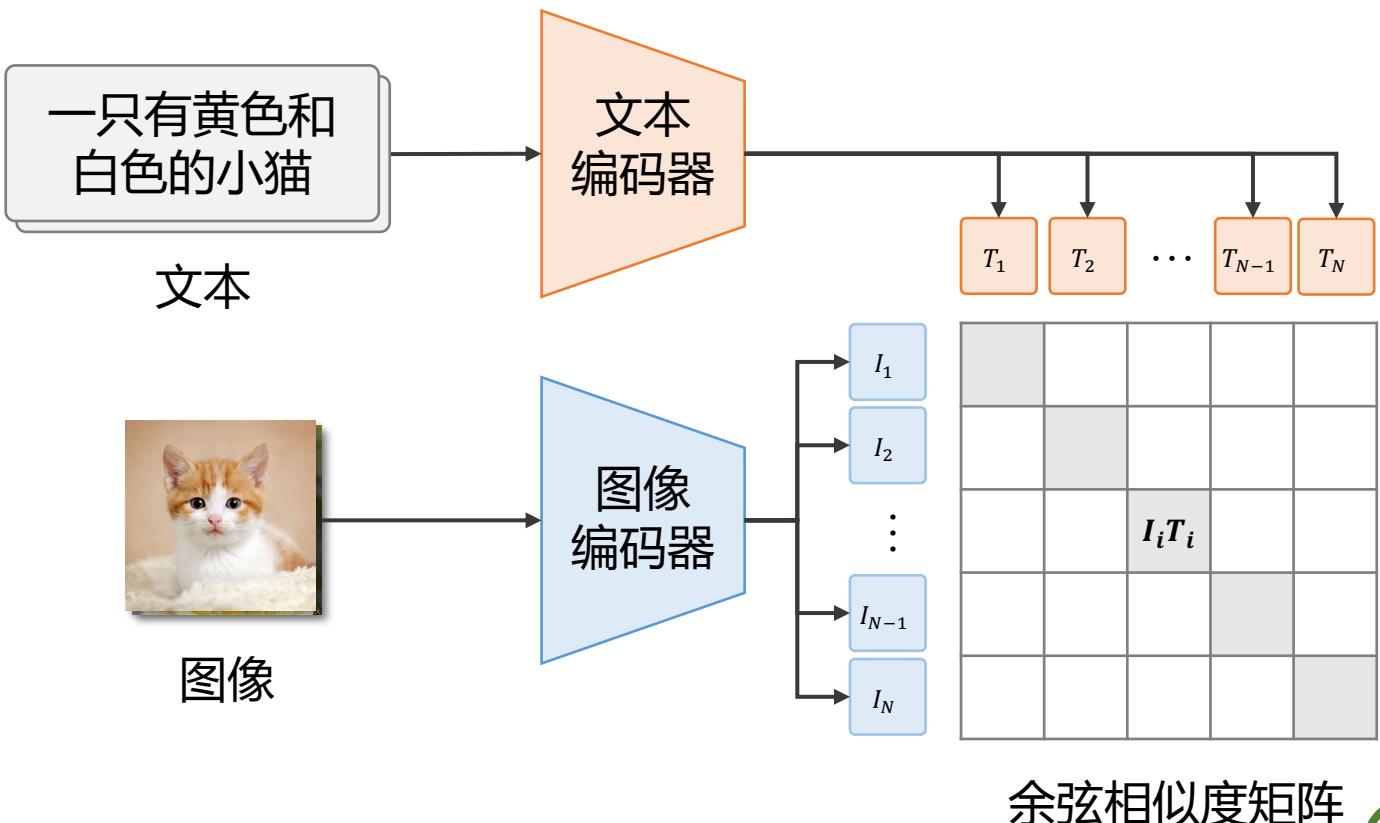


- CLIP模型采用了一种双塔架构，分别是图像编码器和文本编码器，用于处理图像和文本数据。



在文本编码器方面，CLIP沿用了GPT-2的架构；而在图像编码器的设计上，CLIP经过多次尝试，比较了不同的ResNet和Transformer架构，最终选择了ViT模型作为其图像编码器。

CLIP——多模态大模型的先驱



$$S = (I_E \cdot T_E^T) \times e^\varepsilon \in \mathbb{R}^{N \times N}$$

对比
学习

最大化对角线上的相似度值（正确匹配）
尽可能减小非对角线上的值（错误匹配）

- CLIP模型采用了一种双塔架构，分别是图像编码器和文本编码器，用于处理图像和文本数据。
- 采用了自监督学习的方法，通过对大量无标签的图像-文本对进行训练来学习图像和文本之间的对齐关系。
- 通过对比学习的方法，将图像和文本映射到同一个高维语义空间中，使得它们能够在这个空间中相互理解和交互。

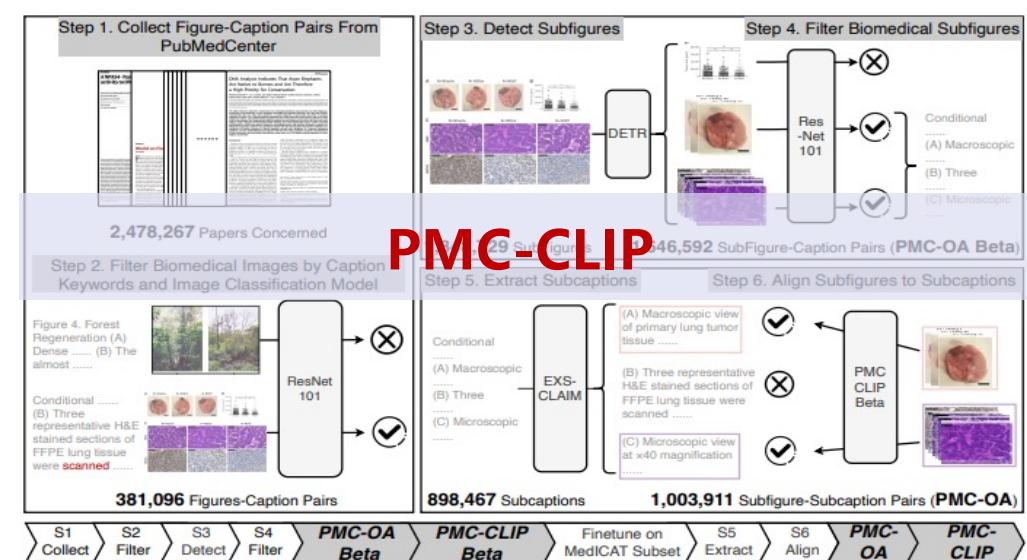
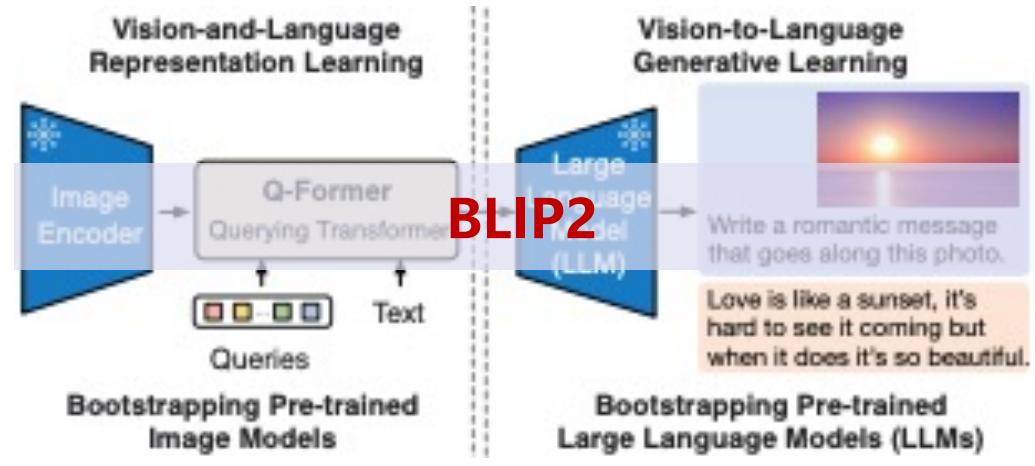
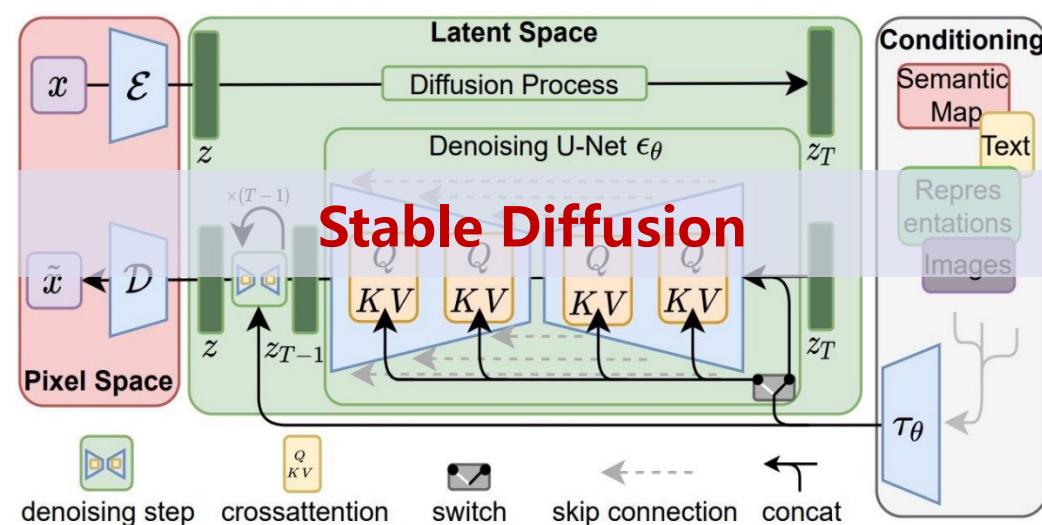
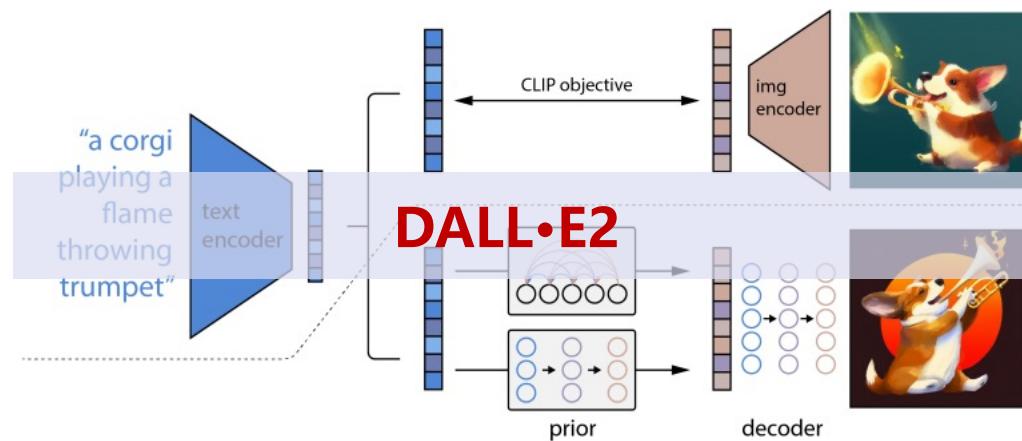
对比学习是一种自监督学习方法，通过将相似样本的特征拉近、不相似样本的特征推远，来学习数据的有效表示。具体来说，模型会构建正样本对（相似样本）和负样本对（不相似样本），通过优化损失函数，使正样本对在特征空间中更接近，负样本对更远离。



对比
学习

CLIP——多模态大模型的先驱

■ CLIP通过对比学习的方式学习到了图像与文本之间的潜在关联，实现了扩模态数据的有效融合，并开创自监督预训练典范，提升迁移学习能力。



知识点3：从基础大模型到垂域大模型

学
习
无
所
不
及



01 基础大模型核心要点

02 垂域大模型的构建与应用

■ 基础大模型是基石，是底座，是基座。



■ 垂域大模型是在特定领域精耕细作的“尖兵”。



■ 基础大模型是基于**大规模通用数据**预训练的模型，具备通用语言理解与生成能力。

- 从规模上看，它拥有海量参数，常达数十亿甚至数万亿，像 GPT-3 就有 1750 亿参数，如此规模能捕捉数据中复杂模式。
- 在通用性上，它不是针对特定任务，而是可处理多种类型任务和数据，比如文本、图像、语音等，像 BERT 模型能用于文本分类、问答等多种自然语言处理任务。
- 从功能上，它为下游任务和领域模型提供基础，通过微调适配具体应用场景，减少重新训练成本。



■ 基础大模型采用Transformer架构，依赖海量数据处理与预训练技术，利用分布式训练策略在大规模集群上进行训练。

01

模型架构

基础大模型采用Transformer架构，基于注意力机制，适合并行化训练，提升训练效率。



02

数据处理与预训练

基础大模型对海量数据进行清洗与去噪，去除无用或错误信息。



03

模型训练与优化

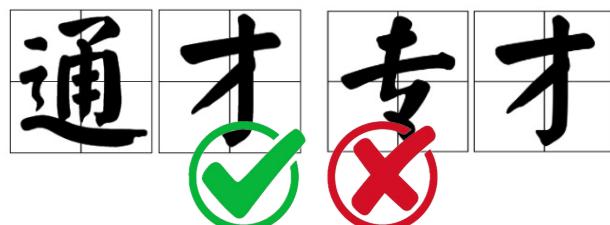
基础大模型采用数据并行和模型并行，加速模型训练过程。



■ 基础大模型虽然通用性强，但是专业性不足，并且存在数据安全与隐私问题，以及脏/假数据对于模型的损害风险，同时面临计算资源与推理效率挑战。

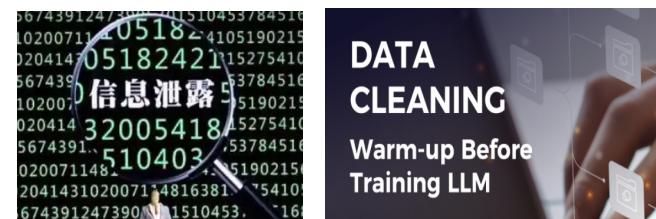
通用性与专业性

基础大模型通用性强，但在专业领域表现不够精准，难以满足精细化需求。在处理专业术语和复杂任务时存在局限，难以满足特定领域的专业需求。



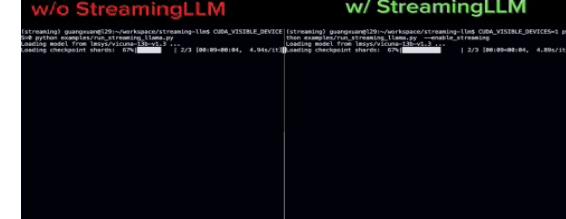
数据安全与隐私

基础大模型使用大量通用数据，涉及数据安全和隐私问题。需加强数据加密和隐私保护，确保数据安全，数据安全是基础大模型的重要问题。



计算资源与效率

基础大模型训练和推理需要大量计算资源，计算成本较高。计算资源是基础大模型的瓶颈，优化资源使用可以提升模型训练和推理效率。



■ 2025年DeepSeek-V3以**低成本、高性能**登顶基础大模型排行榜，其在多个领域的基准测试中表现出色，为开源模型的发展提供了新的强大示例，推动了整个基础大模型应用的技术进步和爆发。

指标	DeepSeek-V3	GPT-4o
训练成本	557.6万美元	约1亿美元
参数量	671亿	1750亿
训练时长	278.8万GPU/小时 (H800)	约2500万GPU/小时 (A100) 
训练硬件	2048块英伟达H800 GPU 	约2.5万个英伟达A100 GPU 
训练效率	高效负载均衡、FP8混合精度训练 	未公开具体优化技术
性价比	显著高于GPT-4o	较低

Benchmark (Metric)	DeepSeek-V3	Qwen2.5 72B-Inst.	Llama3.1 405B-Inst.	Claude-3.5- Sonnet-1022	GPT-4o 0513
Architecture	MoE	Dense	Dense	-	-
# Activated Params	37B	72B	405B	-	-
# Total Params	671B	72B	405B	-	-
MMLU (EM)	88.5	85.3	88.6	88.3	87.2
MMLU-Redux (EM)	89.1	85.6	86.2	88.9	88
MMLU-Pro (EM)	75.9	71.6	73.3	78	72.6
DROP (3-shot F1)	91.6	76.7	88.7	88.3	83.7
English	IF-Eval (Prompt Strict)	86.1	84.1	86	86.5
GPQA-Diamond (Pass@1)	59.1	49	51.1	65	49.9
SimpleQA (Correct)	24.9	9.1	17.1	28.4	38.2
FRAMES (Acc.)	73.3	69.8	70	72.5	80.5
LongBench v2 (Acc.)	48.7	39.4	36.1	41	48.1
HumanEval-Mul (Pass@1)	82.6	77.3	77.2	81.7	80.5
LiveCodeBench(Pass@1-COT)	40.5	31.1	28.4	36.3	33.4
LiveCodeBench (Pass@1)	37.6	28.7	30.1	32.8	34.2
Code	Codeforces (Percentile)	51.6	24.8	25.3	20.3
SWE Verified (Resolved)	42	23.8	24.5	50.8	38.8
Aider-Edit (Acc.)	79.7	65.4	63.9	84.2	72.9
Aider-Polyglot (Acc.)	49.6	7.6	5.8	45.3	16
Math	AIME 2024 (Pass@1)	39.2	23.3	23.3	16
	MATH-500 (EM)	90.2	80	73.8	78.3
	CNMO 2024 (Pass@1)	43.2	15.9	6.8	13.1
Chinese	CLUEWSC (EM)	90.9	91.4	84.7	85.4
	C-Eval (EM)	86.5	86.1	61.5	76.7
	C-SimpleQA (Correct)	64.1	48.4	50.4	51.3

✓ 高性能 ✓ 低成本 ✓ 高训练效率 ✓ 开源

- 垂域大模型，也被称为领域大模型或行业大模型，是相对于通用基础大模型而言，针对特定垂直领域或行业进行专门设计、训练和优化的大型人工智能模型。
- 设计原则可遵循三阶段优化路径：领域数据收集与标注、知识融合以及模型优化。

领域针对性强

数据专业性强

任务特定性强

知识深度强

01

领域数据收集与标注：从专业数据库、行业报告、企业内部数据等多渠道收集数据，制定严格的标注规范，确保数据标注准确性和一致性。

02

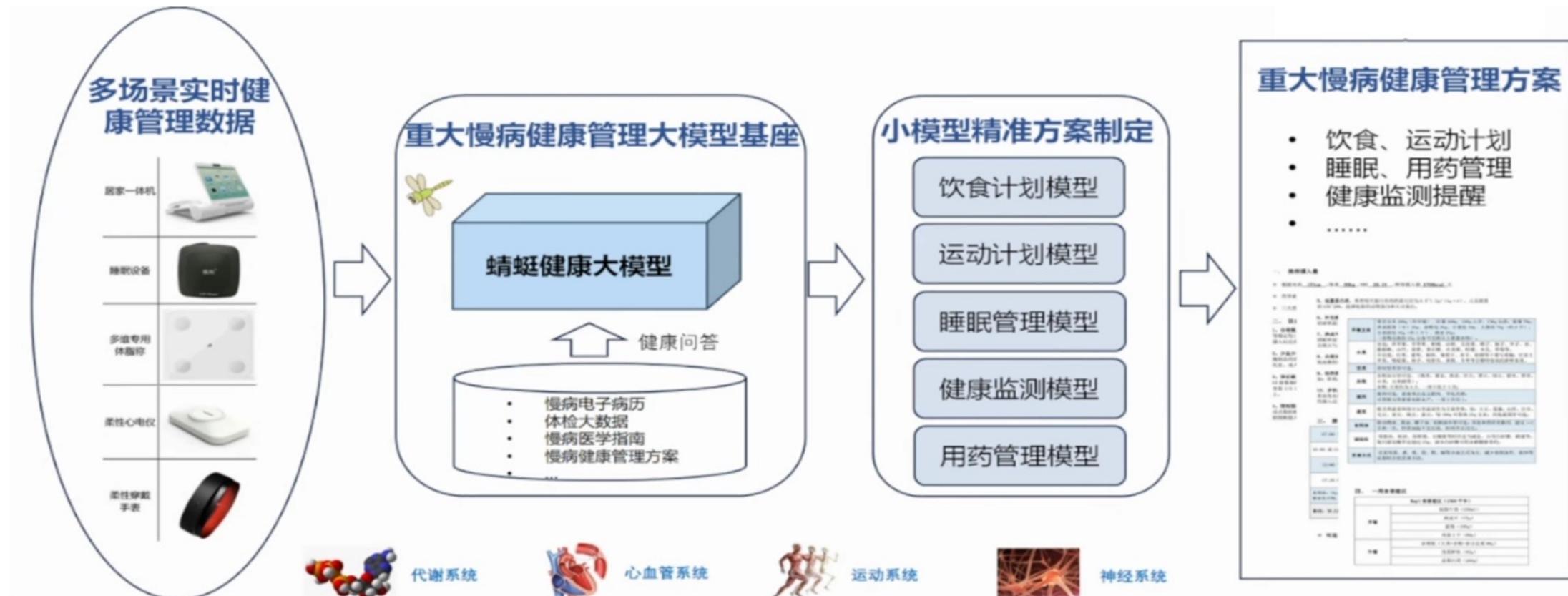
知识融合：构建领域知识图谱，将领域知识融入模型训练中，通过知识蒸馏、多模态融合等技术，提升模型对领域知识的理解和应用能力。

03

模型优化：根据领域数据特点和任务需求，对模型架构进行调整和优化，如增加特定领域的模块、调整网络层数和参数，以提高模型性能。

垂域大模型——你想要的专才！

- 垂域大模型，也被称为领域大模型或行业大模型，是相对于通用基础大模型而言，针对特定垂直领域或行业进行专门设计、训练和优化的大型人工智能模型。
- 设计原则可遵循三阶段优化路径：领域数据收集与标注、知识融合以及模型优化。

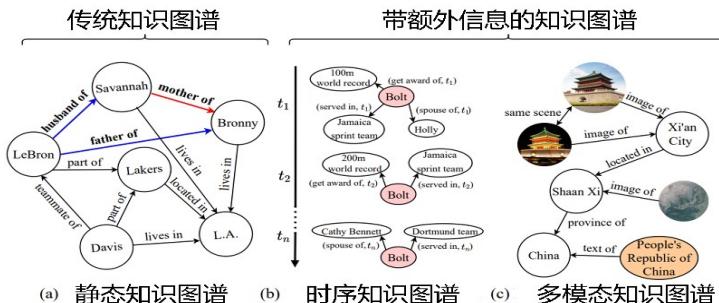


■ **领域知识深度融合技术**进一步推动垂域大模型发展。



知识图谱增强推理

方法：实体链接+图神经网络，通过实体链接将文本实体与知识图谱对应，再用图神经网络推理实体间关系。



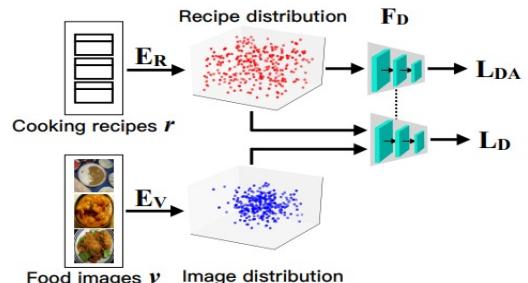
领域规则引导解码

方法：受限解码+规则模板，
依据领域规则限制解码过程，
确保生成内容符合规则。



多模态数据对齐

方法：跨模态对比学习，通过对比学习使不同模态数据特征对齐，提升模型对多模态数据理解。

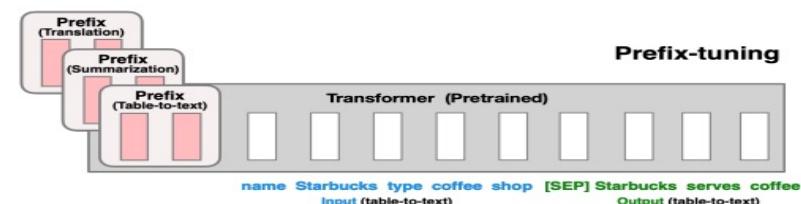
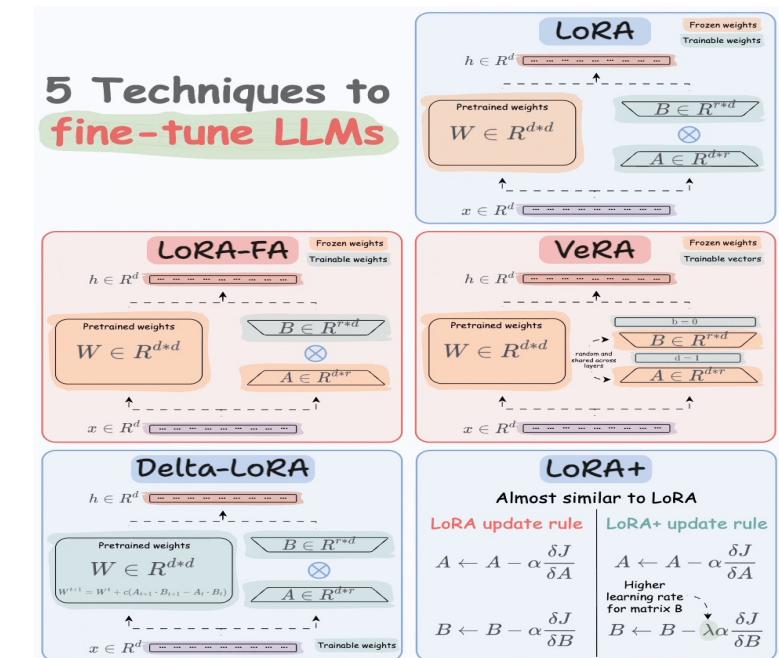


■ **参数高效微调技术**是垂域大模型实现高效定制化的技术手段。

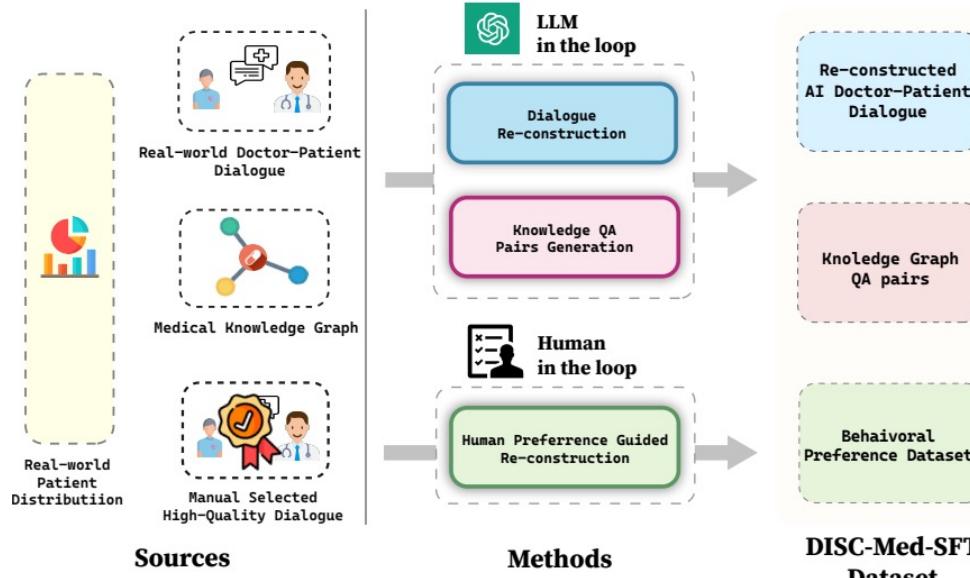
全量 Supervised Fine-tuning: 使用大量标注数据进行微调，使模型在特定领域任务上达到更好的性能。参数量100%，训练成本高，适用于数据充足、算力充裕场景，能充分优化模型性能。

LoRA: 通过低秩矩阵分解进行微调。在保持原模型性能的同时，适应特定领域的任务需求。参数量0.1-1%，训练成本极低，适用于低资源垂域适配，能在少量数据下快速优化模型。

Prefix-tuning: 通过调整模型的提示词，引导模型输出。参数量0.1%，训练成本中等，适用于生成式任务控制，可有效控制生成内容方向。



垂域大模型的构建与应用



医疗大模型DISC-MedLLM



安恒信息恒脑大模型

基于**Baichuan13BBase**基础大模型指令微调得到

DISC-Med-SFT训练集，包含超过**47万个**来源于现有的多个医疗数据集的不同示例。

以**大模型为底座，以智能体为核心**，实现安全能力场景化调度

2024年唯一入选工信部“未来产业领域”案例名单的安全垂域大模型

知识点4：“AI”上创作

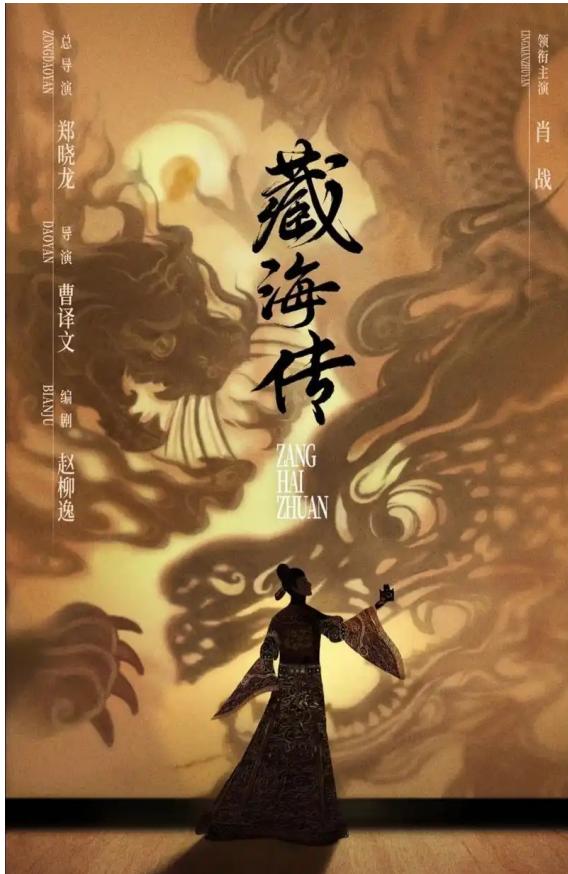
学
习
有
深
度



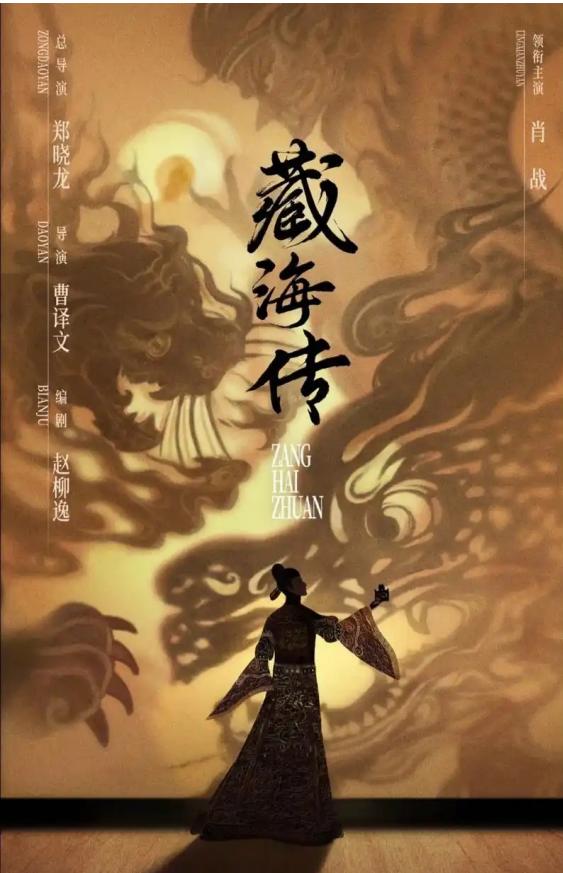
01 生成式人工智能

02 生成式设计与艺术创作

猜一猜，哪个是AI生成的？



猜一猜，哪个是AI生成的？



■ **生成式人工智能 (Generative Artificial Intelligence, GAI)** 的提出源于机器学习领域对于创造性任务的探索，它标志着人工智能从简单的数据分析和模式识别，向能够自主生成新内容的能力迈进。



《郴州客居感怀》
霓虹隐现九寰深，义帝陵前草木暗。
轨接湘南云外驿，楼吞岭表日边金。
纷纭世相屏中演，寂寞乡愁酒后沉。
莫问郴江流昼夜，车潮涨落即潮音。

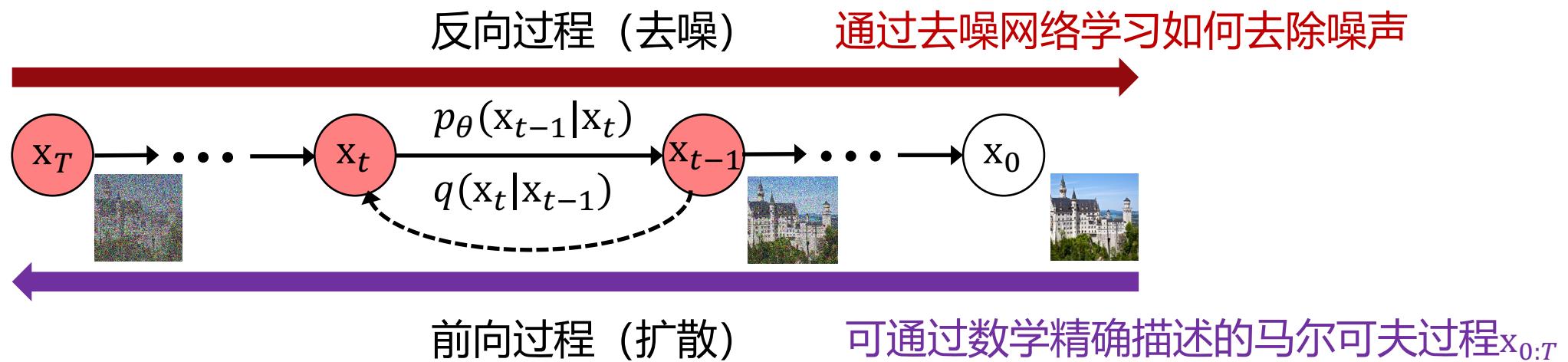




生成式人工智能指能够生成新的数据或内容的人工智能技术，更强调技术层面，专注于如何通过算法生成高质量的合成数据或内容。

人工智能生成内容指通过人工智能技术生成的内容。内容形式可以包括文本、图像、音频、视频等。目标是生成具有创意和独特性的内容。

■ **扩散模型 (Diffusion Models, DM)** 基于非平衡热力学原理，定义了一个马尔可夫链，将数据分布逐步转化为噪声分布，再通过学习逆过程从噪声中恢复数据，实现了高质量的数据生成，成为当前生成模型领域的重要支柱之一。



前向扩散过程 (加噪过程) : 模型不断对输入数据加入噪声，直到其变成纯高斯噪声

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t}x_{t-1}, \beta_t I)$$

反向采样过程 (去噪过程) : 模型则逐步去除噪声并恢复出原始数据。

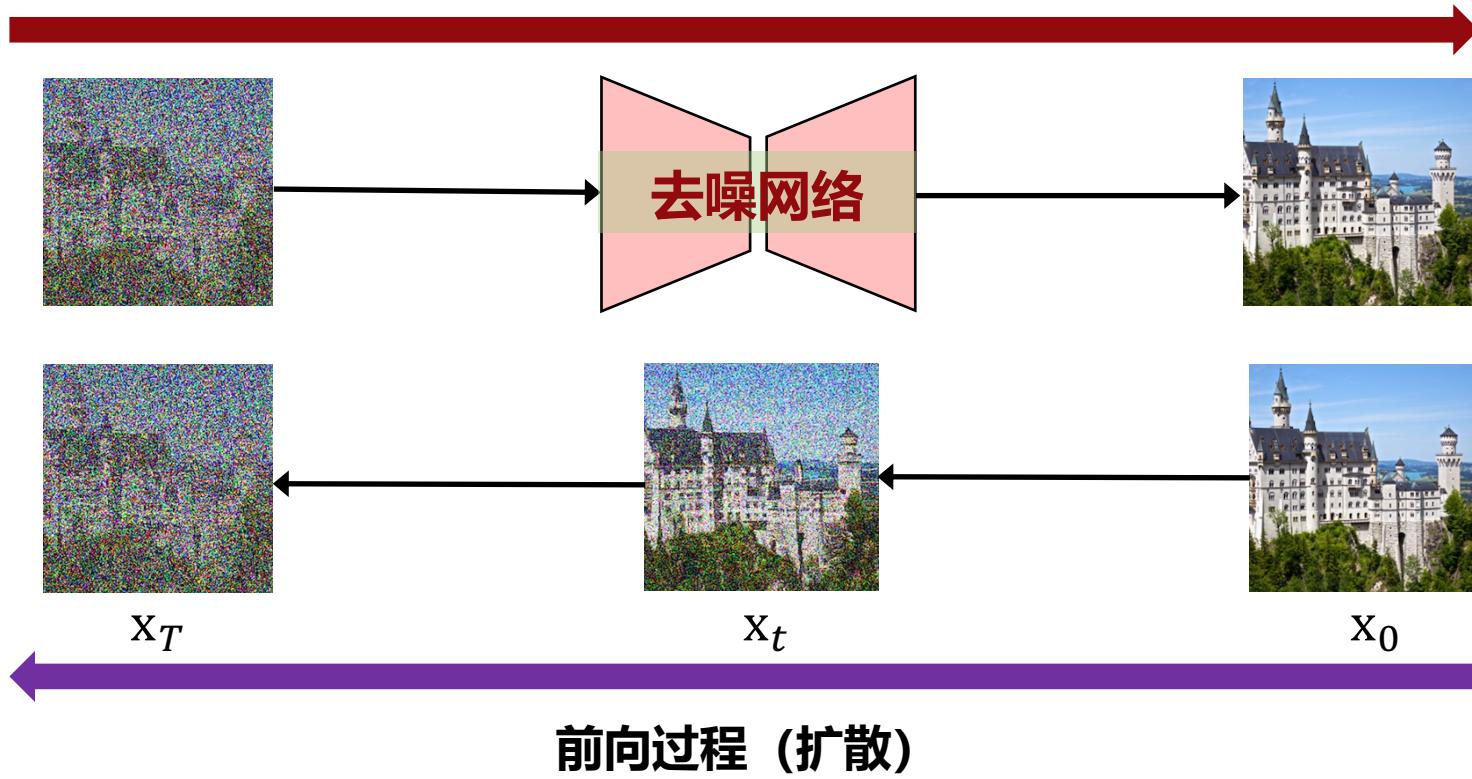
$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \sigma_\theta^2(x_t, t)I)$$

生成式人工智能关键技术——扩散模型

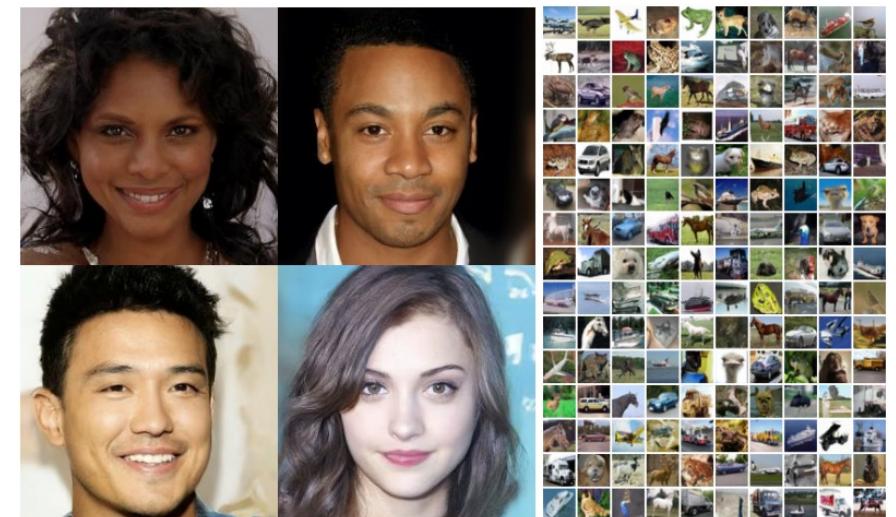
扩散模型中去噪网络训练目标是最小化 x_T 与标准高斯噪声直接的差异：

$$\mathbb{E}_{t, \mathbf{x}_0, \epsilon} [\|\epsilon - \epsilon_\theta(\sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon, t)\|^2]$$

反向过程（去噪）



DDPM生成图片效果



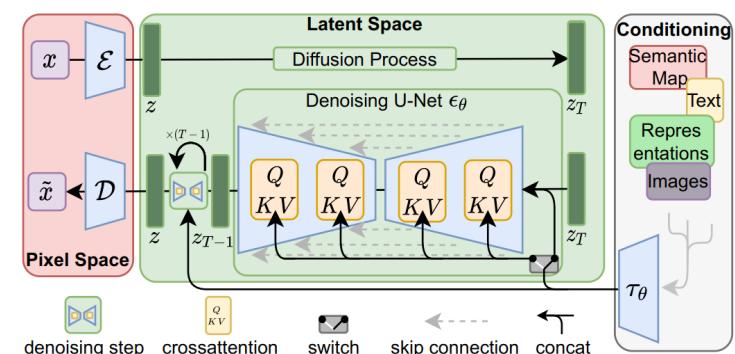
■ 隐式扩散模型 (**Latent Diffusion Models, LDM**) 在低维隐空间进行扩散，证明了 CLIP 等文本-图像对齐模型在生成任务中的有效性，首次将高性能扩散模型完全开源，定义生成式 AI 新标准。

DDPM

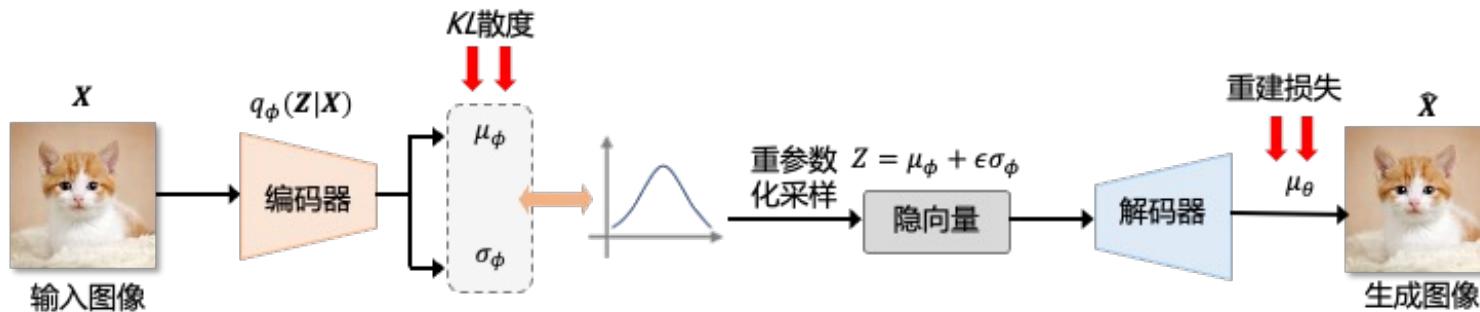
- 在原始图像上进行加噪和去噪的，效率低下
- 图像分辨率很大时，计算成本和时间成本高昂
- 对图像存在大量冗余信息进行扩散影响模型稳定性



LDM: 将原始图像映射到低维隐空间中编码向量，对其进行扩散和采样操作，**大大减少了计算量和内存占用**。同时，在潜在空间中数据分布更简单、更规则，模型**更容易学习和捕捉**数据的特征与分布规律，建模难度降低。



生成式人工智能关键技术图谱



变分自编码器 (VAE)

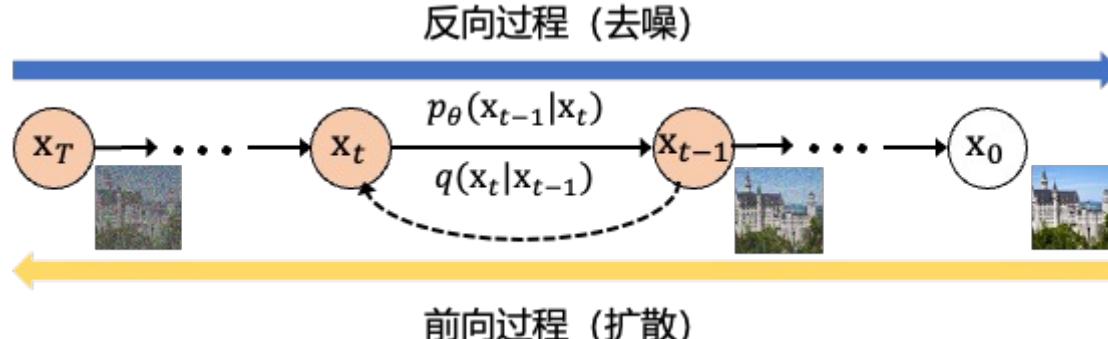
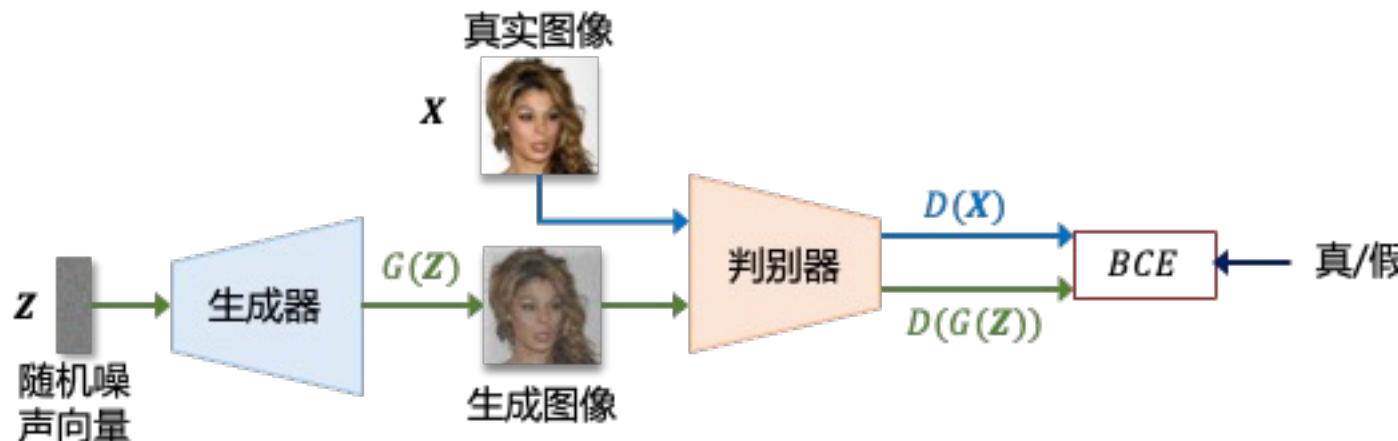
- 可解释性较强
- 训练过程涉及复杂的数学计算
- 生成样本质量较低

生成对抗网络 (GAN)

- 生成样本质量较高
- 训练过程不稳定
- 可解释性不强

扩散模型 (DM)

- 生成样本效果逼真
- 可解释性强
- 训练过程慢，成本高昂



生成式设计与艺术创作

The screenshot shows the Stable Diffusion online interface. At the top left is a grid of generated images. In the center is a large image of a girl with curly hair holding a book, styled like a Rembrandt painting. On the right is a detailed description of the prompt: "Close-up portrait of a smiling girl holding a book, oil painting in the style of Rembrandt". Below the main image are various editing tools and settings, including a sidebar for styles and a preview area.

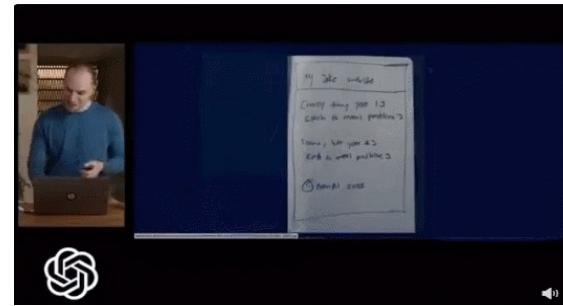
Close-up portrait of a smiling girl holding a book, oil painting in the style of Rembrandt

The screenshot shows a PPT slide titled "生成式设计与艺术创作". The slide features a generated image of a modern skyscraper with a grid pattern. The slide includes a logo, a progress bar indicating "已阅读 30 个网页", and a text input field asking for a presentation about generative design and art. The footer shows a navigation bar with various slide thumbnails.

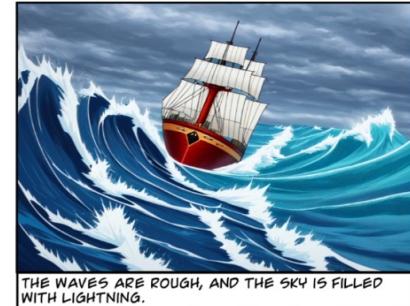
生成式设计与艺术创作



音乐创作



网页制作



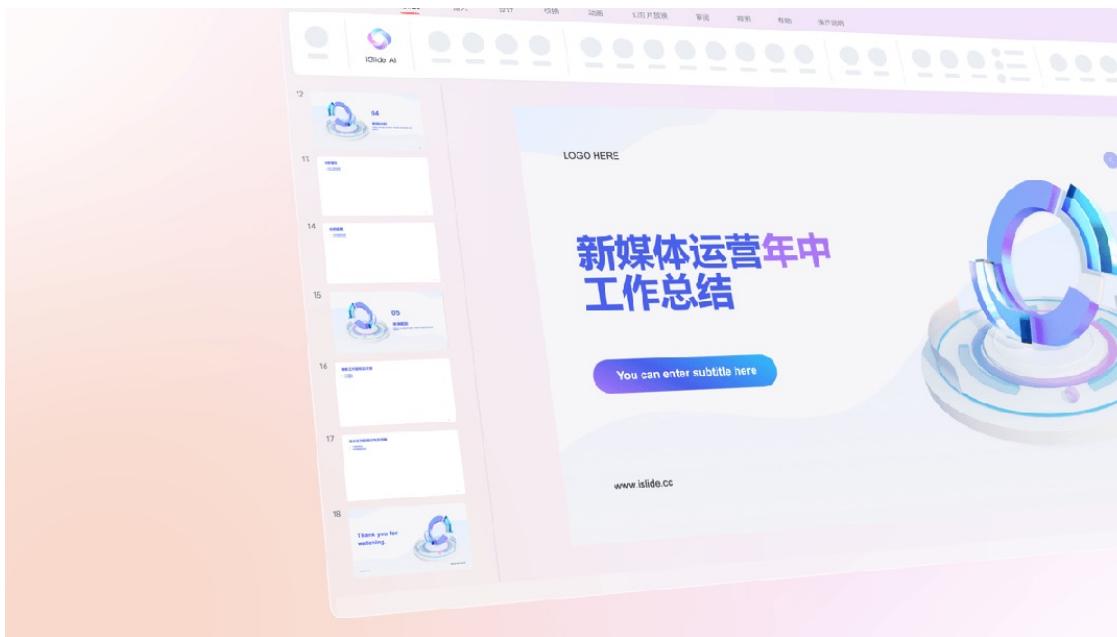
漫画创作



动画生成



剧本创作



人民日报《AI共创大片 | 江山如此多娇》



AI主播连麦京津冀



新华社“AIGC绘中国”



人民网“数字主持人速览两会”

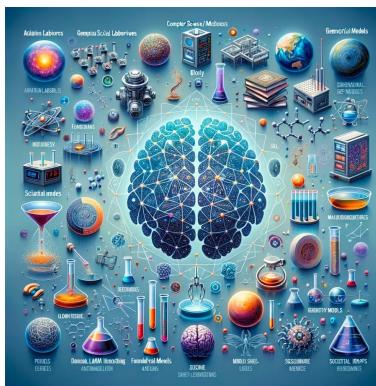


新质生产力首次被写入

央视新闻《AI数“读”两会》

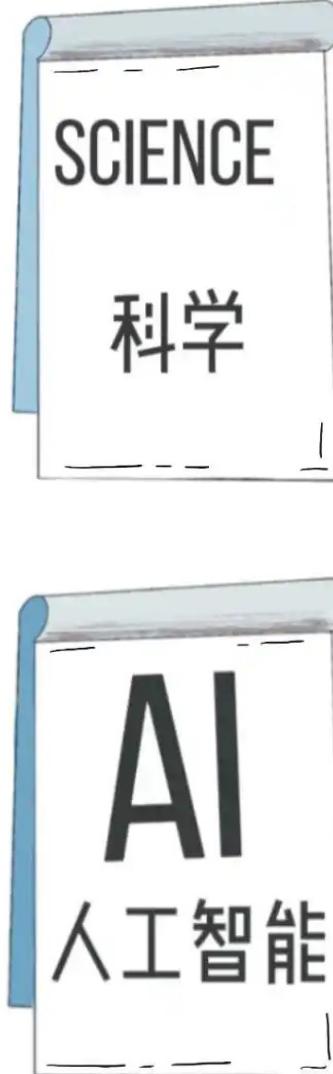
知识点5： 全新的科学发现范式——AI for Science

学
习
有
深
境



01 人工智能与自然科学

02 人工智能与人文社科



AI for Science

AI不仅可以搞科学，
而且搞得很科学。



科学发现新范式——AI for Science



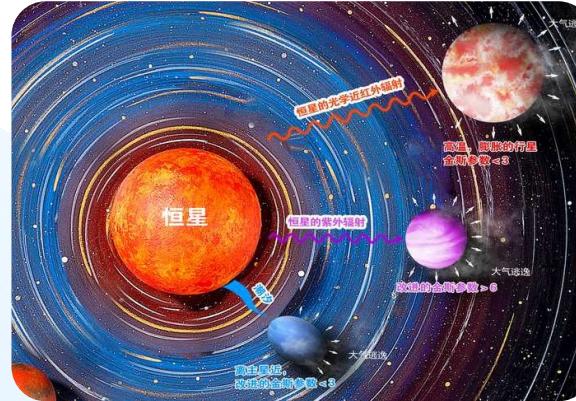
化学



医学



天文

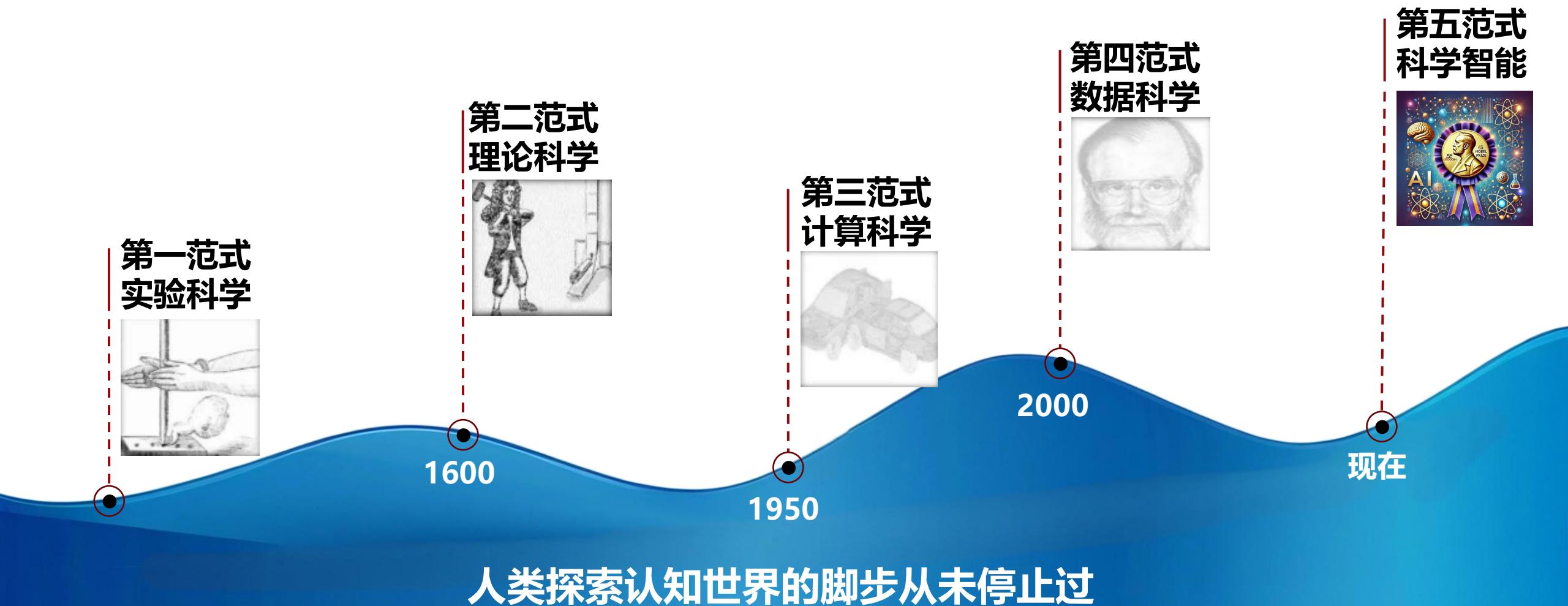


生物



走向未来

- 人工智能技术为人类发现和总结规律提供了新的**第五范式：科学智能/人工智能驱动的科学研究 (AI for Science或AI4Science)**。



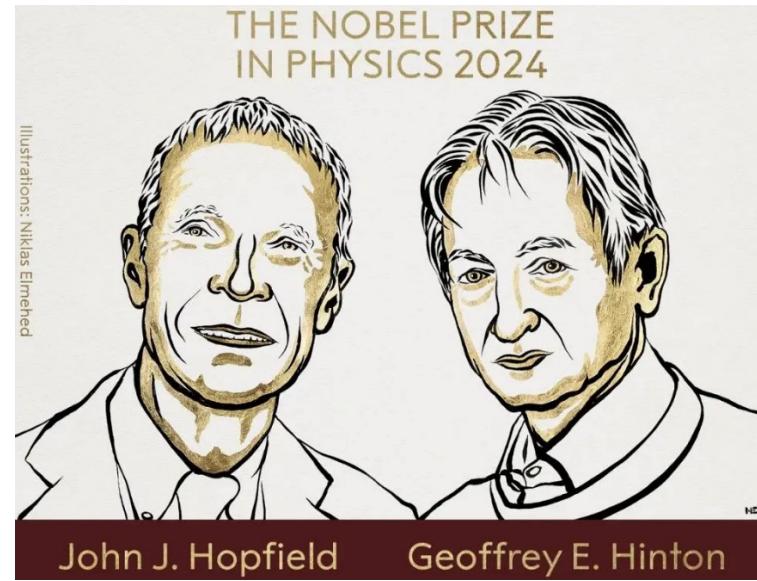
■ 人工智能科学家获得诺贝尔奖标志着科学范式的重大转变，体现了AI在科学研究中的革命性地位及其对传统学科边界的突破，它揭示了数据驱动与机理研究的互补性，以及跨学科融合的必然性。

诺贝尔化学奖



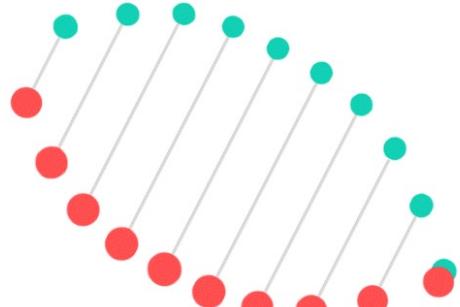
AlphaFold

诺贝尔物理学奖

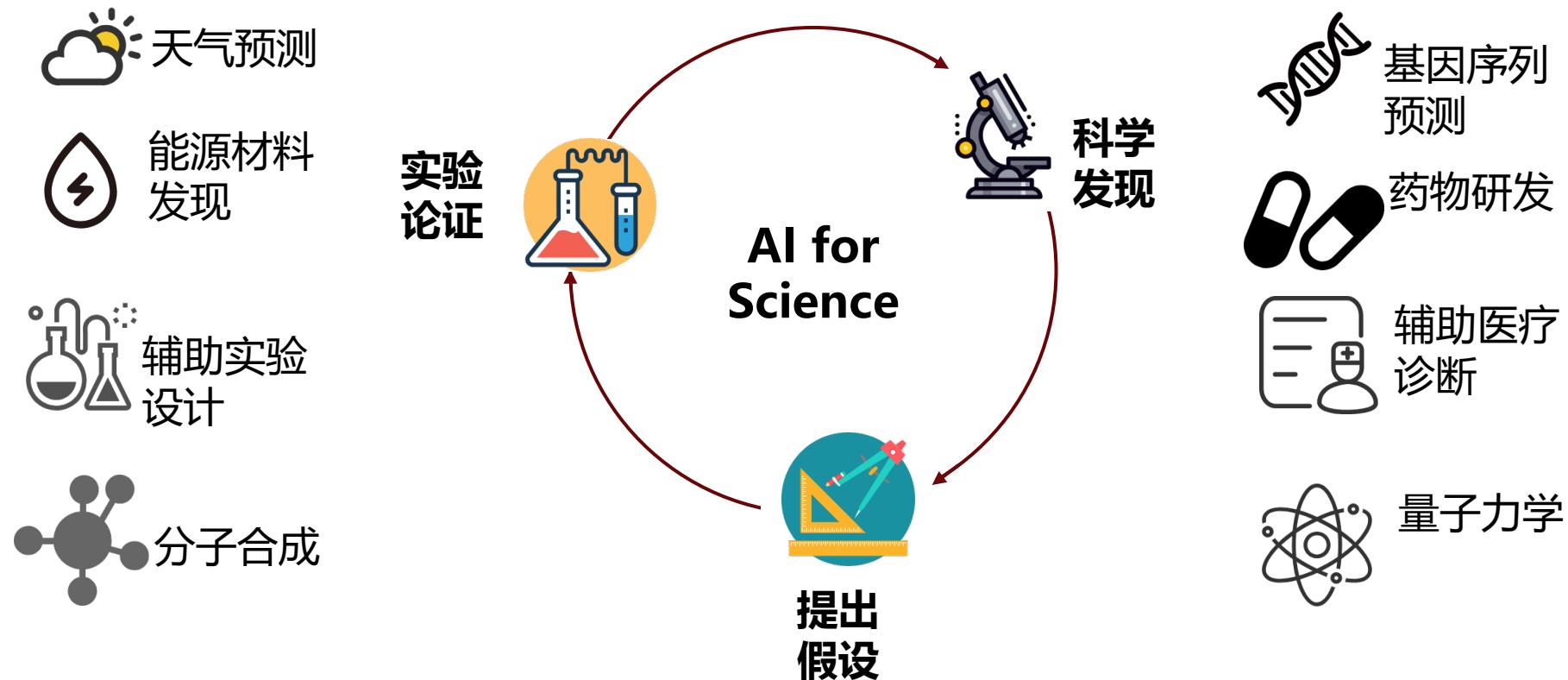


人工神经网络机器学习

AI+科学 “双螺旋引擎”
共振驱动



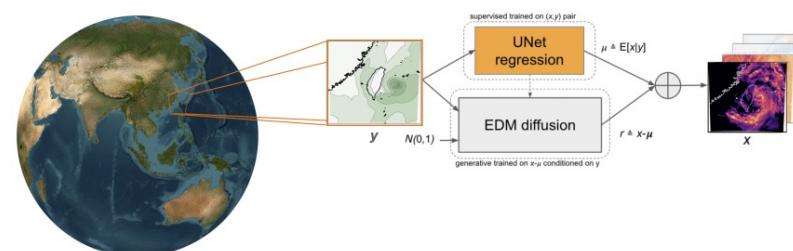
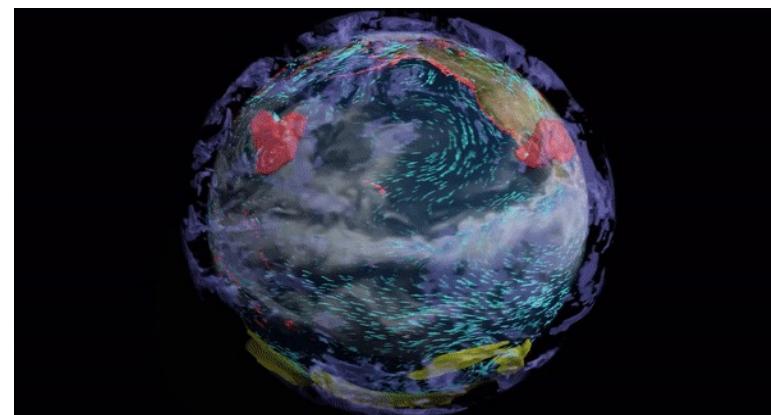
■ AI与自然科学的协同效应正在形成良性循环：AI不仅从自然科学汲取灵感，也开始反过来推动自然科学研究的突破性进展，这种双向促进的典范案例正不断涌现。



■ 人工智能与自然科学的结合正在推动科学研究的范式转变，为物理、化学、生物、地球科学等领域带来了前所未有的机遇和突破。其目标是利用机器学习、深度学习、自然语言处理和其他人工智能技术，加速科学发现，改善实验设计，提高数据分析的效率，并揭示新的科学知识和理论。

传统天气预报

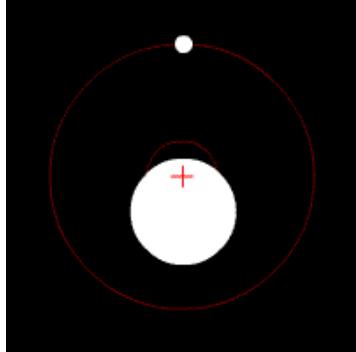
- 基于数值天气预报
- 难以实现公里级精细化模拟
- 长周期预测误差大
- 成熟体系需要数十年验证
- 计算成本高，依赖超算



AI天气预报

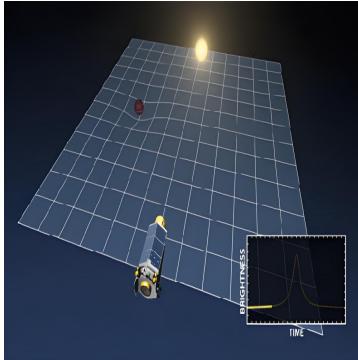
- 数据驱动与物理规律融合
- 公里级细节
- 1分钟内完成传统超算数小时的7天全球预报
- 极端事件精准预警
- 可生成数千种可能天气场景，辅助灾害风险评估

■ 人工智能在天文学、材料科学以及生物学等领域已经出现了实际落地的应用。



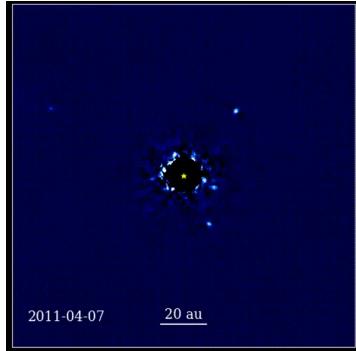
天体测量法

- 需要高仪器精度
- 对行星要求高
- 需数十年观测数据积累
- 数据处理复杂



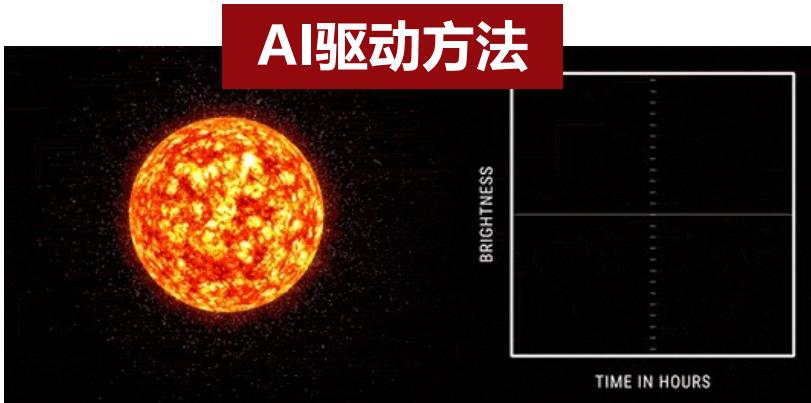
微引力透镜法

- 事件罕见且短暂
- 难以重复观测
- 对宿主恒星信息有限



直接成像法

- 目前仅能探测距离恒星较远、大质量的行星
- 难以捕捉类地行星或宜居带内天体



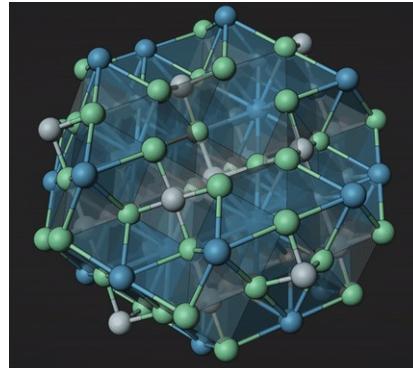
AI驱动方法

快速处理和分析海量天文数据
精准识别行星信号
探索行星系统演化



谷歌和NASA利用“猎星代码”从开普勒-90系统中发现了第八颗行星—开普勒-90i，以及从开普勒-80系统中发现了最小行星—开普勒-80g

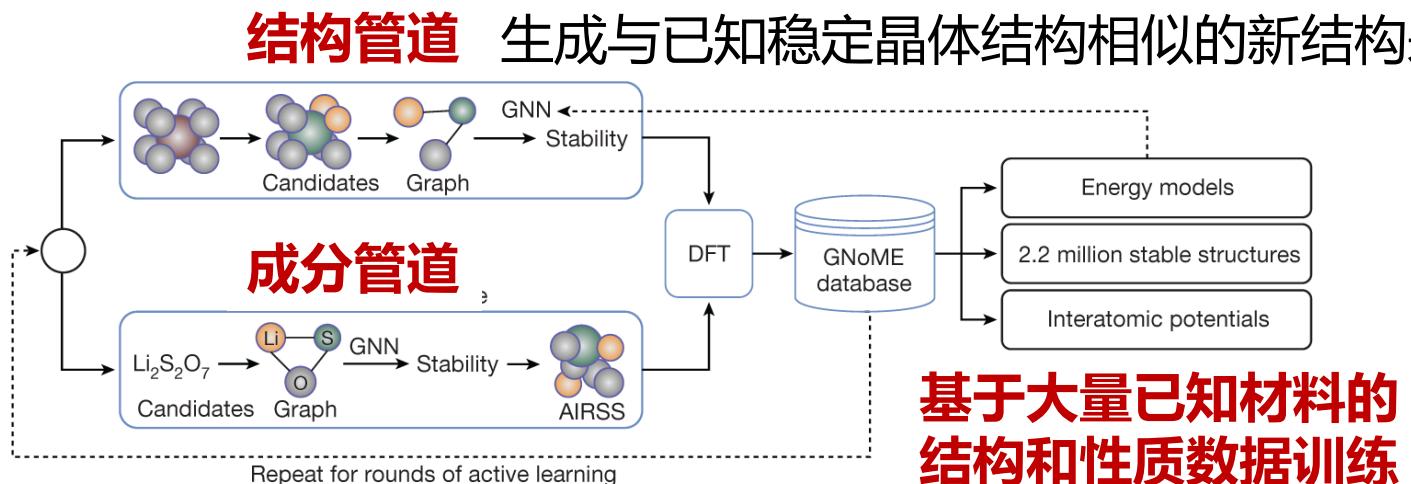
■ 人工智能在天文学、材料科学以及生物学等领域已经出现了实际落地的应用。



传统方法依赖于繁琐的实验和复杂的数据分析，周期长、效率低。

传统方法需要大量的实验设备和材料，成本较高。

传统方法往往受到现有理论和经验的限制，难以发现突破性的新材料。

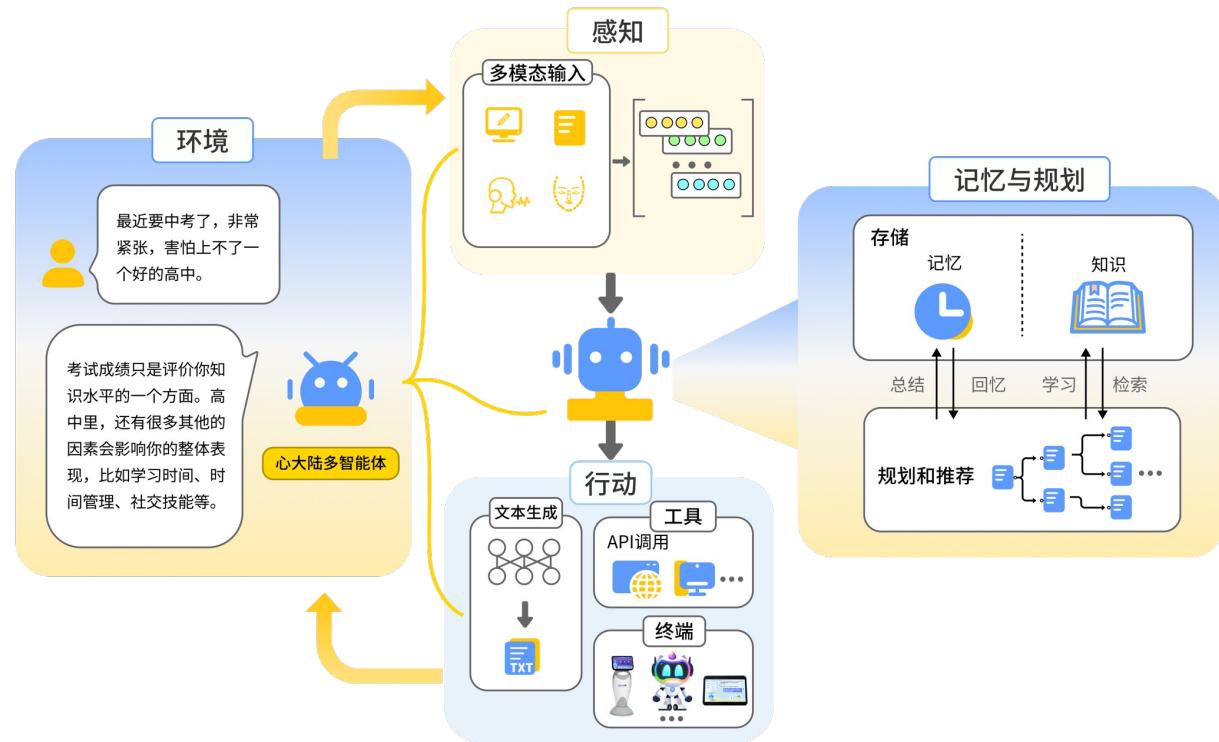


已预测 220 万种新材料结构

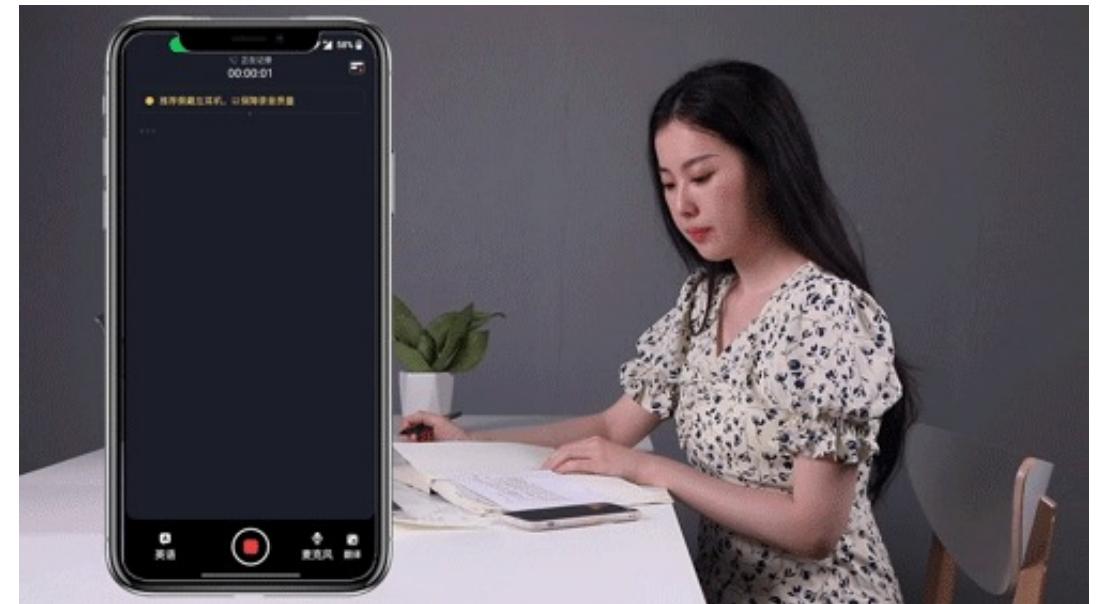
已创造 700 多种新材料

遵循基于化学式的更随机的方法，通过随机组合不同的化学元素来生成新的材料候选者。

■ 语言学通过对语言的结构、功能、演变等方面的研究，为理解人类文化和社会提供重要基础。自然语言处理助力语言学数字化转型，提升语言理解与应用效率，推动人文科学领域的研究和创新，加速知识传播与文化发展。

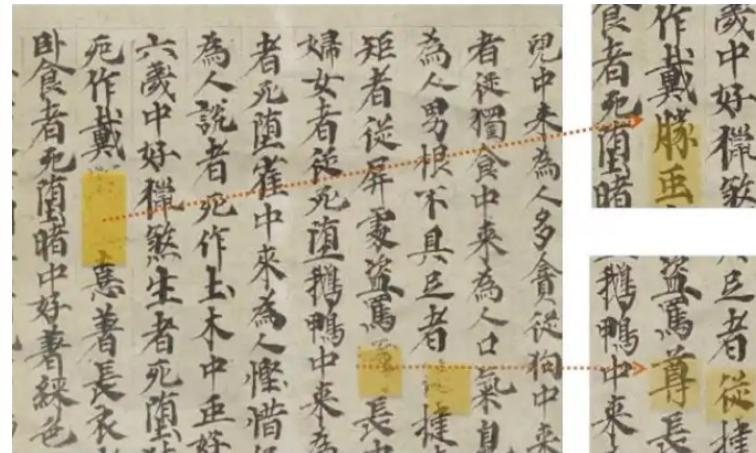


AI情感分析与陪伴



AI同声传译

■ 通过三维扫描技术获取文物的数字模型，然后利用人工智能算法进行三维重建，能够高精度地还原文物的原始状态，为破损文物修复决策提供依据。



“敦煌遗书”修复



“永乐宫壁画”修复

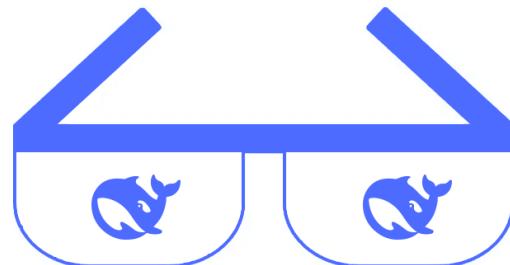


AI虚拟修复三星堆文物



AI还原文物细节纹理痕迹

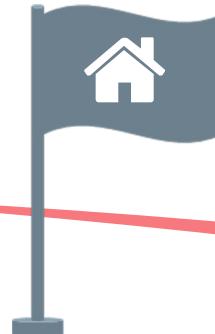
知识点6：玩转大模型



01 大模型与AIGC

02 玩转大模型

学
习
有
进
境



—2021

LLM理论与技术准备



2022

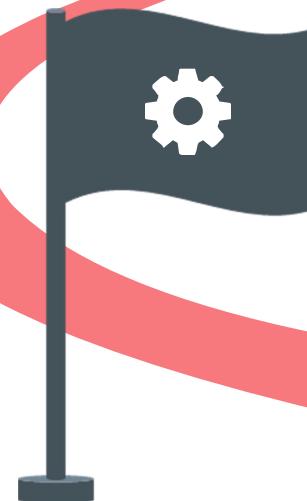


ChatGPT引爆LLM研发热潮



2023

LLM飞速发展，
全民AI浪潮涌现



2024



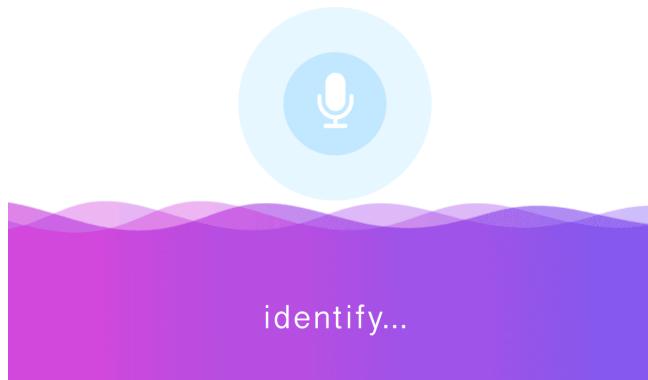
基于LLM的Agent
百花齐放，垂域大
模型称为焦点



2025



DeepSeek释放
LLM深度思考能力



智能语音



智慧城市



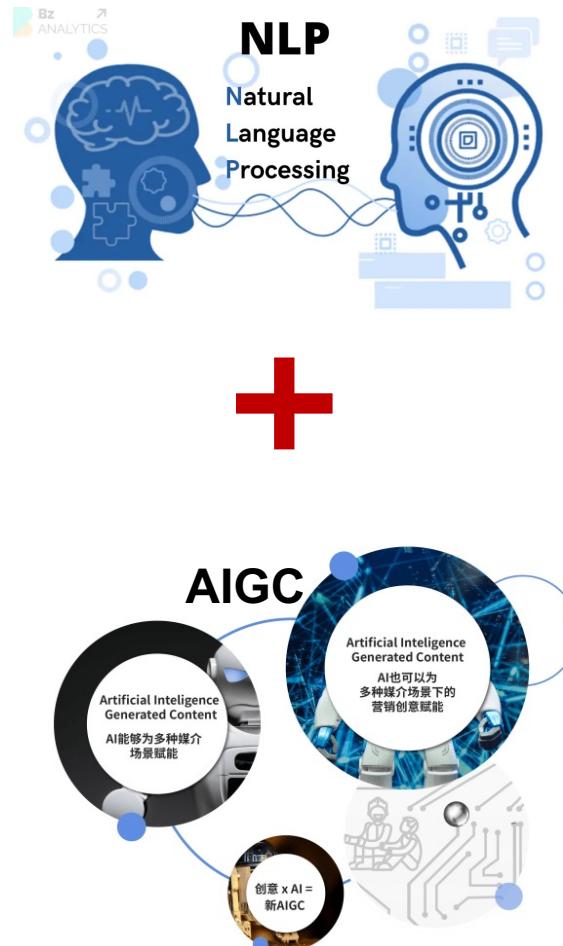
自动驾驶



智能办公



- 人工智能生成内容 (AI-Generated Content, AIGC) 是人工智能进入全新发展时期的重要标志，其核心思想是利用人工智能算法生成具有一定创意和质量的内容。



用户体验升级
行业创新与转型

内容创作领域革新
生产力提升与成本降低

OpenAI



ChatGPT

百度



文心一言

阿里巴巴



通义千问

科大讯飞



星火大模型

智谱AI



智谱清言

月之暗面



Kimi

字节跳动



豆包大模型

深度求索



deepseek

文本生成

语言理解

知识问答

逻辑推理

+

专业应用软件

||

智能、高效、个性化内容生成

提示词



向AI表达需求

引导生成

提高精度

增强交互

简洁清晰的指令

必要的任务背景

明确的任务期望

用户向大模型输入的文本内容，用于触发大模型的响应并指导其如何生成或回应

■ 文本类应用实践——文档撰写

案例

请帮我写一份面向国际听众的演讲稿。需要以**中国南北方豆腐脑的甜咸口味差异**为案例，讲解不同地域的**饮食文化特色**，并延伸到**中西方餐饮习惯的对比**。要求**语言生动有趣，适合跨文化背景的观众理解**。



中国饮食文化差异之趣谈 —— 以豆腐脑为例

尊敬的各位国际友人、亲爱的国内朋友们：

大家好！今天，非常荣幸能站在这里，和大家分享一些关于中国饮食文化的有趣差异，我们将会从一场别具一格的“舌尖之争”——南北方甜咸豆腐脑之战切入，进而探索更广阔的饮食文化天地，也希望能为大家了解中西方饮食文化差异提供一个独特视角。

一、豆腐脑“甜咸大战”：南北方的味蕾分歧

豆腐脑，这道以大豆为原料制作而成、口感细腻嫩滑的美食，在中国大江南北都广受欢迎。然而，当涉及到它的“最佳搭档”——是甜口还是咸口时，南北方的朋友们可就“各执一词”了。

在南方，甜豆腐脑堪称经典。清晨的街边早餐店，一碗热气腾腾的豆腐脑端上桌，上面淋满了晶莹剔透的红糖糖浆，有的还会撒上一把炒得香脆的花生碎，或是点缀几粒软糯的桂花蜜饯。那红糖的香甜、花生的酥脆、桂花的清香，与细嫩的豆腐脑完美融合，一口下去，仿佛整个味蕾都被这甜蜜的滋味温柔包裹，开启活力满满的一天。对于南方人来说，这甜豆腐脑不仅仅是一份早餐，更是一种从小到大的味觉记忆，是家乡的味道，是温暖与甜蜜的象征。

而北方的咸豆腐脑则别有一番风味。卤汁是咸豆腐脑的灵魂，通常由黄花菜、木耳、香菇、肉末等食材精心熬制而成，色泽浓郁，香气扑鼻。吃的时候，舀上一勺吸饱了卤汁的豆腐脑放入口中，鲜咸的味道瞬间充盈口腔，丰富的配料更是增添了多层次的口感，既有蔬菜的清新，又有肉末的醇厚，让人回味无穷。北方人偏爱这咸口，认为它滋味醇厚，能与豆腐脑的清淡形成鲜明对比，更能抵御北方的寒冷气候，带来一种扎实的满足感。

三、从豆腐脑之争看中西方饮食文化差异

当我们把视野从豆腐脑的甜咸之争拓展到中西方饮食文化时，会发现更多有趣的差异与共性。

在食材选择上，中国饮食讲究食材的多样性与季节性，追求“不时不食”，无论是天上飞的、地上跑的，还是水里游的，都能巧妙地变成餐桌上的美味佳肴。西方饮食则相对更注重食材的分类与标准化，对肉类、蔬菜、谷物等有明确的划分与搭配比例，更倾向于使用特定的几种食材进行烹饪，强调食材的品质与来源。

烹饪技法方面，中国烹饪技法繁多且精湛，煎、炒、烹、炸、蒸、煮、炖、烤等，每一种技法都能赋予食材独特的风味与口感，厨师们凭借经验和感觉，灵活运用火候、调料等元素，创造出千变万化的菜肴。西方烹饪则更注重科学与规范，通过精准的温度控制、时间把握以及标准化的调料配方，确保每一道菜肴的品质稳定，像烘焙面包、制作牛排等都有严格的烹饪流程与标准。

饮食观念上，中国人注重饮食的平衡与养生，讲究五味调和、荤素搭配，认为饮食与身体健康息息相关，餐桌上的每一道菜都蕴含着对健康与生活的美好祝愿。西方人则更强调饮食的营养均衡，关注食物中的蛋白质、碳水化合物、脂肪、维生素等营养成分的摄入比例，将饮食视为维持生命体征与健康的重要手段。

■ 文本类应用实践——文档撰写



我要写一篇关于**当前国内外主流通用大模型的文档介绍**，受众为**高校的大学生、研究生和企业研究人员**，需要**通俗易懂、体现一定的专业性**，并**以表格形式总结这些大模型的功能、优缺点**，内容控制在**3000字**左右（DeepSeek+WPS灵犀）

三、技术趋势与挑战

1. 发展趋势

- 多模态融合：如商汤“商量SenseChat”支持视觉、语音、文本协同交互 ①。
- 端云协同：端侧算力提升，降低云端依赖并增强隐私保护 ②。
- 开源与闭源并行：开源模型（如Llama2、DeepSeek）推动技术民主化，闭源模型（如GPT-4）保障商业竞争力 ② ⑨。

2. 核心挑战

- 算力瓶颈：训练万亿参数模型需高成本投入，国产芯片生态仍需突破 ② ⑤。
- 数据安全与隐私：模型可能泄露敏感信息或生成误导性内容 ④ ⑩。
- 可解释性不足：黑盒特性限制其在医疗、法律等高风险领域的应用 ④

四、应用场景与行业落地



领域	典型案例	代表模型
金融	风险评估、投资决策辅助（通义千问、智谱清言）	① ③
医疗	疾病诊断辅助、医学文献分析（DeepSeek、Kimi）	⑧ ⑩

二、国内外主流大模型概览

1. 国内大模型

模型名称	所属公司/机构	核心功能	技术特点	优点	缺点	典型应用场景
文心一言	百度	知识问答、创意生成、多模态处理	知识增强型，跨语言/跨模态能力	复杂提示理解，支持代码调试	特定领域，专业性不足	搜索、教育、智能办公 ① ③
通义千问	阿里云	多轮对话、文案创作、逻辑推理	多模态理解，长文本输入支持	性价比高，灵活性强	复杂任务处理效率待提升	金融、医疗、物流 ① ③
星火	科大讯飞	语音交互、数学推理、代码生成	跨领域知识理解，语音处理能力突出	审核速度快，逻辑推理强	代码生成能力较弱	客服、科研、教育 ① ③
混元	腾讯	开放域聊天、音视频创作	结合搜索引擎，支持多格式文档解析	个性化智能体创建体验佳	专业性领域需优化	内容创作、口语陪练 ① ③
智谱清言	智谱AI	推理与创作、商业分析	中英双语对话，亿万字符预训练	功能全面，适合高复杂度任务	多轮对话，连贯性不足	广告文案、科研辅助 ① ③

已生成文档

一文读懂大模型：从基础到实战与未来

大模型基础入门

重点解释大模型的完整定义，包括广义和狭义概念，次要阐述其作为神经网络模型的本质特征，普通段落精选 300 字

一文读懂大模型：从基础到实战与未来

一、大模型基础入门

WPS 灵犀

大模型这个词如能只是听说过，并不完全清楚它到底是什么。其实，大模型的世界丰富多彩，从它的定义、本质，到分类以及和其他相关概念的区别，都有着独特的内涵。接下来，就让我们一起深入探索大模型基础入门的知识，揭开它神秘的面纱。

1. 大模型的定义与本质

大模型的定义，广义上来说，是指具有大规模参数和强大计算能力的人工智能模型。它就像是一个知识渊博的“超级大脑”，能够处理和分析海量的数据。而狭义的大模型，则更侧重于那些在自然语言处理、计算机视觉等领域展现出卓越性能的模型。

■ 文本类应用实践——文档撰写



我要写一篇关于**当前国内外主流通用大模型的文档介绍**，受众为**高校的大学生、研究生和企业研究人员**，需要**通俗易懂、体现一定的专业性**，并**以表格形式总结这些大模型的功能、优缺点**，内容控制在**3000字**左右（DeepSeek+豆包）

一、大模型基础入门

1. 大模型的定义与本质

- 配图位置：章节标题下方
- 提示词：

"透明玻璃质感的3D大脑模型，内部填充二进制代码与数学公式，神经网络线条包裹地球仪，背景悬浮GPU芯片与数据流，赛博朋克蓝紫光效，8K细节"

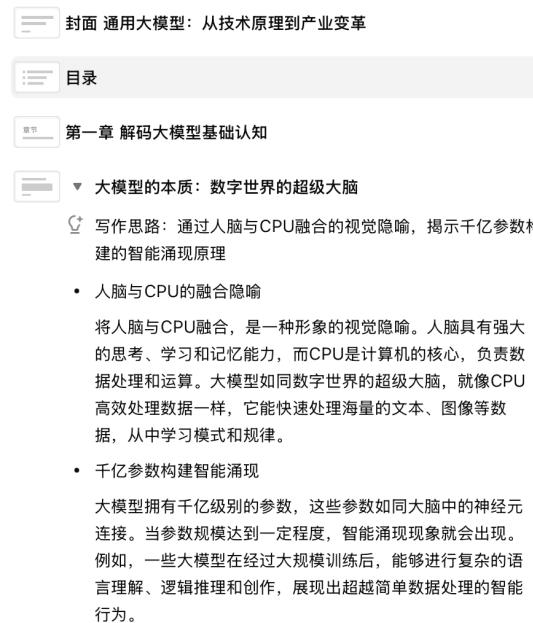


考虑到公众号文章的推广，内容的吸引力，我需要**在对应的位置加入图片丰富内容**，请给出建议，**并且对插入的图片生成合适的提示词用于生成图片**

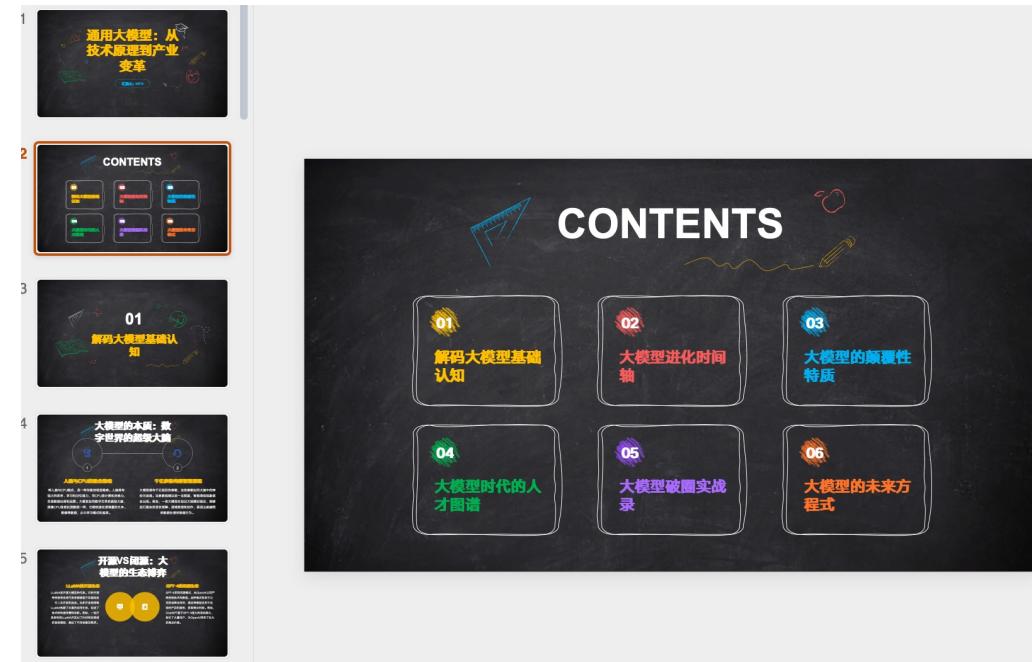
■ 文本类应用实践——PPT制作



我要写一篇关于**当前国内外主流通用大模型的文档介绍**，受众为**高校的大学生、研究生和企业研究人员**，需要**通俗易懂、体现一定的专业性**，并**以表格形式总结这些大模型的功能、优缺点**，内容控制在**3000字**左右（DeepSeek+WPS灵犀）



封面 通用大模型：从技术原理到产业变革
目录
章节 第一章 解码大模型基础认知
大模型的本质：数字世界的超级大脑
写作思路：通过人脑与CPU融合的视觉隐喻，揭示千亿参数构建的智能涌现原理
人脑与CPU的融合隐喻
将人脑与CPU融合，是一种形象的视觉隐喻。人脑具有强大的思考、学习和记忆能力，而CPU是计算机的核心，负责数据处理和运算。大模型如同数字世界的超级大脑，就像CPU高效处理数据一样，它能快速处理海量的文字、图像等数据，从中学习模式和规律。
千亿参数构建智能涌现
大模型拥有千亿级别的参数，这些参数如同大脑中的神经元连接。当参数规模达到一定程度，智能涌现现象就会出现。例如，一些大模型在经过大规模训练后，能够进行复杂的语言理解、逻辑推理和创作，展现出超越简单数据处理的智能行为。



CONTENTS

01. 解码大模型基础认知
02. 大模型进化时间轴
03. 大模型的颠覆性特质
04. 大模型时代的人才图谱
05. 大模型破圈实战录
06. 大模型的未来方程式

根据文档内容制作一份PPT，40页左右，并列出大纲

■ 图片类实践——图片生成

我要做一个关于当前国内外主流通用大模型的讲座，受众为高校的大学生、研究生和企业研究人员，需要设计一张海报，请帮我生成对应的提示词



哪吒慵懒斜倚真皮沙发，左手高擎琥珀色珍珠奶茶，右手挥洒出金黄油亮的薯片，腮帮微鼓嘴角沾着星星点点的零食碎屑，赤红混天绫缠绕在电竞椅背。背景中赛博朋克风的现代办公室与传统神话元素碰撞.....



海报呈现深邃的星空蓝背景，渐变交织着流动的量子紫光晕，中央悬浮着立体发光的神经网络结构，金色数据粒子如星轨般穿梭其中.....



参考图



■ 视频类实践——视频生成

案例

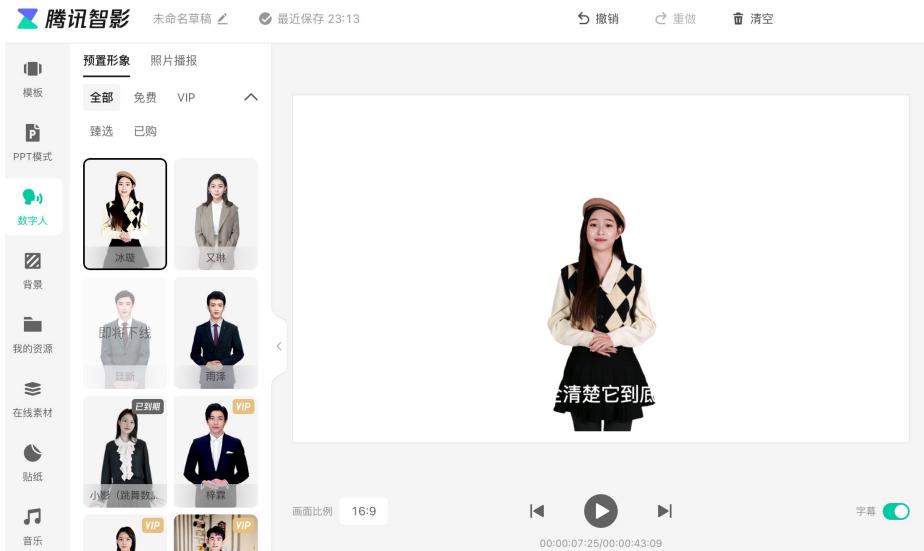
哪吒慵懒斜倚真皮沙发，左手高擎琥珀色珍珠奶茶（奶盖将倾未倾），右手挥洒出金黄油亮的薯片，腮帮微鼓嘴角沾着星星点点的零食碎屑，赤红混天绫缠绕在电竞椅背。背景中赛博朋克风的现代办公室与传统神话元素碰撞——全息财务报表在玻璃幕墙跃动，青铜乾坤圈悬于人体工学办公桌，风火轮造型扫地机器人穿梭于满地零食包装间。悬浮的虾条袋与飘飞的文件页交织出动态韵律，奶茶杯壁凝结的水珠折射出菱形光斑，绛红色混天绫在低饱和莫兰迪色调中成为视觉焦点，营造出传统神祇与现代生活交织的魔幻休闲图景。



■ 视频类实践——数字播报



大模型这个词如今在科技领域可是响当当的存在，但很多人可能只是听说过，并不完全清楚它到底是什么。其实，大模型的世界丰富多彩，从它的定义、本质，到分类以及其他相关概念的区别，都有着独特的内涵。接下来，就让我们一起深入探索大模型基础入门的知识，揭开它神秘的面纱。大模型的定义，广义上来说，是指具有大规模参数和强大计算能力的人工智能模型。它就像是一个知识渊博的“超级大脑”，能够处理和分析海量的数据。而狭义的大模型，则更侧重于那些在自然语言处理、计算机视觉等领域展现出卓越性能的模型。



在科技领域

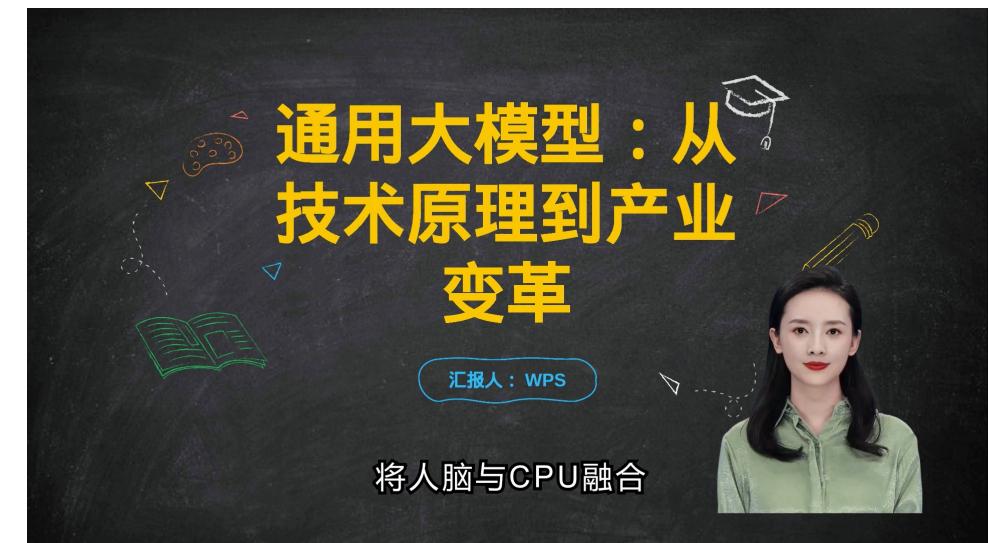
■ 视频类实践——数字播报



上传需要报告的PPT “通用大模型从技术原理到产业变革”

The screenshot shows a digital reporting interface. On the left, there's a sidebar with various tools: 模板 (Template), PPT模式 (PPT Mode), 数字人 (Digital Person), 背景 (Background), 我的资源 (My Resources), 在线素材 (Online Materials), and 贴纸 (Stickers). The main area has a button to "Upload PPT or PDF". Below it, three preview thumbnails are shown:

- Thumbnail 1: A woman in a green shirt with the title "通用大模型：从技术原理到产业变革".
- Thumbnail 2: A woman in a green shirt with the title "CC ITS".
- Thumbnail 3: A woman in a green shirt with the title "解密确认".

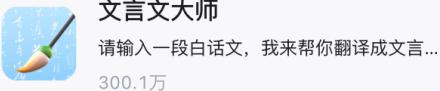


探索 我的智能体

搜索

+ 创建智能体

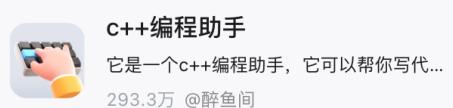
精选 职场 创作 学习 编程 生活 趣味 营销 情感 公文 脑暴 三更多



文言文大师

请输入一段白话文，我来帮你翻译成文言...

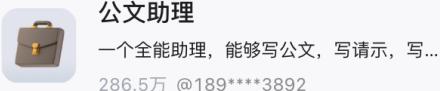
300.1万



C++编程助手

它是一个c++编程助手，它可以帮你写代...

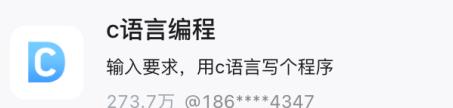
293.3万 @醉鱼间



公文助理

一个全能助理，能够写公文，写请示，写...

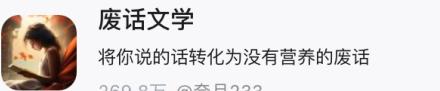
286.5万 @189****3892



c语言编程

输入要求，用c语言写个程序

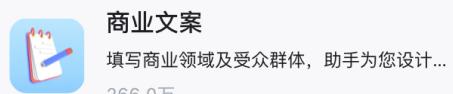
273.7万 @186****4347



废话文学

将你说的话转化为没有营养的废话

269.8万 @奈月233



商业文案

填写商业领域及受众群体，助手为您设计...

266.0万



海龟汤游戏玩家

输入开始游戏，我会帮你出一个具有一定...

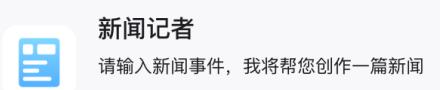
260.5万 @清悦岚



Excel公式编辑器

请输入你的Excel需求场景，我将给你对...

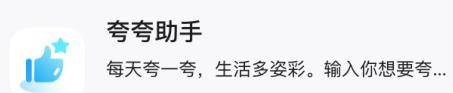
248.2万



新闻记者

请输入新闻事件，我将帮您创作一篇新闻

246.5万 @致Great



夸夸助手

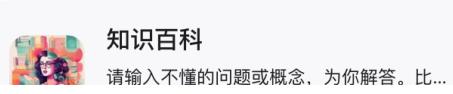
每天夸一夸，生活多姿彩。输入你想要夸...

244.3万 @维莉亚



Python编程助手

回答针对 Python 编码过程中遇见的常见...



知识百科

请输入不懂的问题或概念，为你解答。比...

DeepSeek 实用集成 awesome

将 DeepSeek 大模型能力轻松接入各类软件。访问 [DeepSeek 开放平台](#) 来获取您的 API key。

[English](#)/[简体中文](#)/[日本語](#)

应用程序

	DeepChat	DeepChat 是一款完全免费的桌面端智能助手，内置强大的 DeepSeek 大模型，支持多轮对话、联网搜索、文件上传、知识库等多种功能。
	Wechat-Bot	基于 wechaty 实现的微信机器人，结合了 DeepSeek 和其他 Ai 服务。
	Quantalogic	QuantaLogic 是一个 ReAct（推理和行动）框架，用于构建高级 AI 代理。
	Chatbox	一个支持多种流行LLM模型的桌面客户端，可在 Windows、Mac 和 Linux 上使用
	ChatGPT-Next-Web	一键获取跨平台ChatGPT网页用户界面，支持流行的LLM
	Coco AI	Coco AI 是一个完全开源、跨平台的统一搜索与效率工具，能够连接并搜索多种数据源，包括应用程序、文件、谷歌网盘、Notion、语雀、Hugo 等本地与云端数据。通过接入 DeepSeek 等大模型，Coco AI 实现了智能化的个人知识库管理，注重隐私，支持私有部署，帮助用户快速、智能地访问信息。
	留白记事	留白让你直接在微信上使用 DeepSeek 管理你的笔记、任务、日程和待办清单！

<https://github.com/deepseek-ai/awesome-deepseek-integration>



山东大学
SHANDONG UNIVERSITY

《人工智能通识》AI For Everyone

人工智能前沿

学无止境 气有浩然

教育部-华为“智能基座”课程