

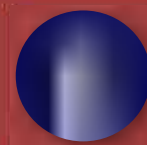
DeepSeek内部研讨系列

DeepSeek与AIGC应用

AI肖睿团队

(孙萍、周嵘、李娜、张惠军、刘誉)

2025年2月20日



- 北大青鸟人工智能研究院
- 北大计算机学院元宇宙技术研究所
- 北大教育学院学习科学实验室



1. 本次讲座为DeepSeek原理和应用系列研讨的讲座之一，主要介绍DeepSeek的基本概念，以及大模型技术和AIGC工具应用，不需要大家具备专业的AI或IT技术背景。
2. 本次讲座首先分析当前备受瞩目的DeepSeek-R1 的概念、优势和历史地位。然后进一步探讨大模型和AIGC的底层工作机制，旨在帮助读者突破工具应用的局限，理解DeepSeek和AIGC的深层次价值。最后，介绍如何科学选择与高效使用AI工具，为大家提供更具深度与实用性的应用场景的指导，给听众带来更落地的AI应用价值。
3. 尽管 DeepSeek-R1 以其低成本和开源策略为行业带来变革，但当前网络上的大量相关内容仅停留在工具应用层面，易对初级AI应用人员造成概念和思维方式的误导，这也是本次讲座希望解决的问题。



目录

CONTENTS

01 详解DeepSeek R1

02

AIGC的概念和应用

03 AIGC的能力揭秘

04

选择AIGC工具



PART 01 ▶

详解DeepSeek-R1

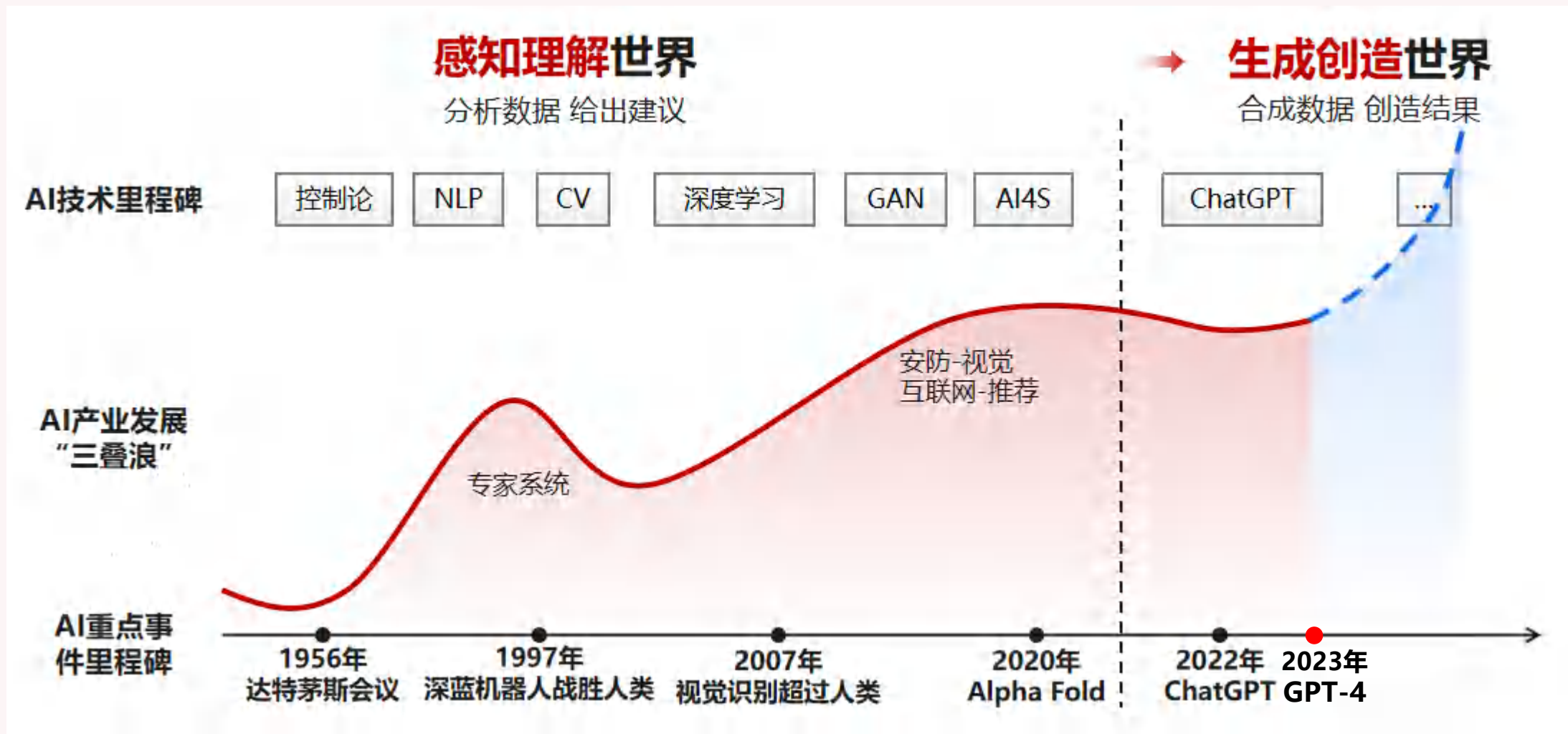
本部分介绍了DeepSeek-R1模型的技术特性、发展历程、应用场景及其在AIGC领域的定位。

首先介绍人工智能的发展历程，以及大模型相关术语，并对比DeepSeek-R1与其他模型的性能表现。DeepSeek-R1以其低成本、开源策略和卓越的推理能力脱颖而出，尤其在复杂逻辑推理、数学和编程任务中表现优异。

随后介绍DeepSeek公司的背景、市场定位以及DeepSeek-R1的技术原理和应用场景，揭示了其在推理密集型任务、教育、科研、知识应用和文档分析等领域的独特优势，并列举接入该模型的第三方应用。

通过对人工智能发展历史以及DeepSeek-R1的介绍和分析，本部分旨在为听众提供一个对DeepSeek的客观、全面的认识，并理解该模型在AIGC领域的重要地位和应用潜力。

人工智能发展历程



AIGC的发展历程



大模型相关术语

- 多模态
 - 文本、图片、音频、视频
- AI工具（国内）
 - DeepSeek、豆包、Kimi、腾讯元宝、智谱清言、通义千问、秘塔搜索、微信搜索...
- 通用模型
 - 大语言模型 (LLM, Large Language Model)
 - 视频模型
 - 多模态模型
 -
- 行业模型（垂直模型、垂类模型）
 - 教育、医疗、金融等



大模型的关键发展

关键进展



生成模型



推理模型

生成模型与推理大模型的对比

比较项	GPT-4o (生成模型)	DeepSeek-R1 (推理模型)
模型定位	专注于通用自然语言处理和多模态能力，适合日常对话、内容生成、翻译以及图文、音频、视频等信息处理、生成、对话等。	侧重于复杂推理与逻辑能力，擅长数学、编程和自然语言推理任务，适合高难度问题求解和专业领域应用，在中文表达上容易出彩。
推理能力	在日常语言任务中表现均衡，但在复杂逻辑推理（如数学题求解）上准确率较低。	在复杂推理任务表现卓越，尤其擅长数学、代码推理任务，在部分基准测试（如 GPQA）中准确率高于 GPT-4o。
多模态支持	支持文本、图像、音频乃至视频输入，可处理多种模态信息。	当前主要支持文本输入，不具备图像处理等多模态能力。
应用场景	适合广泛通用任务，如对话、内容生成、多模态信息处理以及多种语言相互翻译和交流；面向大众市场和商业应用。	适合需要高精度推理和逻辑分析的专业任务，如数学竞赛、编程问题和科学研究；在思路清晰度要求高的场景具有明显优势，比如采访大纲、方案梳理；在对中文语言表达和情感表达方面有明显优势。
用户交互体验	提供流畅的实时对话体验，支持多种输入模态；用户界面友好，适合大众使用。	可展示大部分链式思考过程，便于专业用户理解推理过程；界面和使用体验具有较高的定制性，但整体交互节奏较慢。

推理模型的优劣势

优势	劣势
- 演绎或归纳等推理能力强（如谜题、数学证明）	- 响应速度慢且计算成本高（需要更多推理时间）
- 链式思维推理出色（善于分解多步骤问题）	- 基于知识的任务更容易出错（容易产生幻觉）
- 擅长复杂决策任务	- 处理简单任务的时候效率低（容易"过度思考"）
- 可以呈现思考过程	

常见推理模型

DeepSeek R1

OpenAI o1

OpenAI o3-mini

Gemini 2.0

Grok3

Kimi 1.5

火爆全网的DeepSeek-R1

当地时间1月27日，受中国人工智能初创公司——深度求索公司（DeepSeek）冲击，美国人工智能主题股票遭抛售，美国芯片巨头英伟达（NVIDIA）股价历史性暴跌，纳斯达克综合指数大幅下跌。

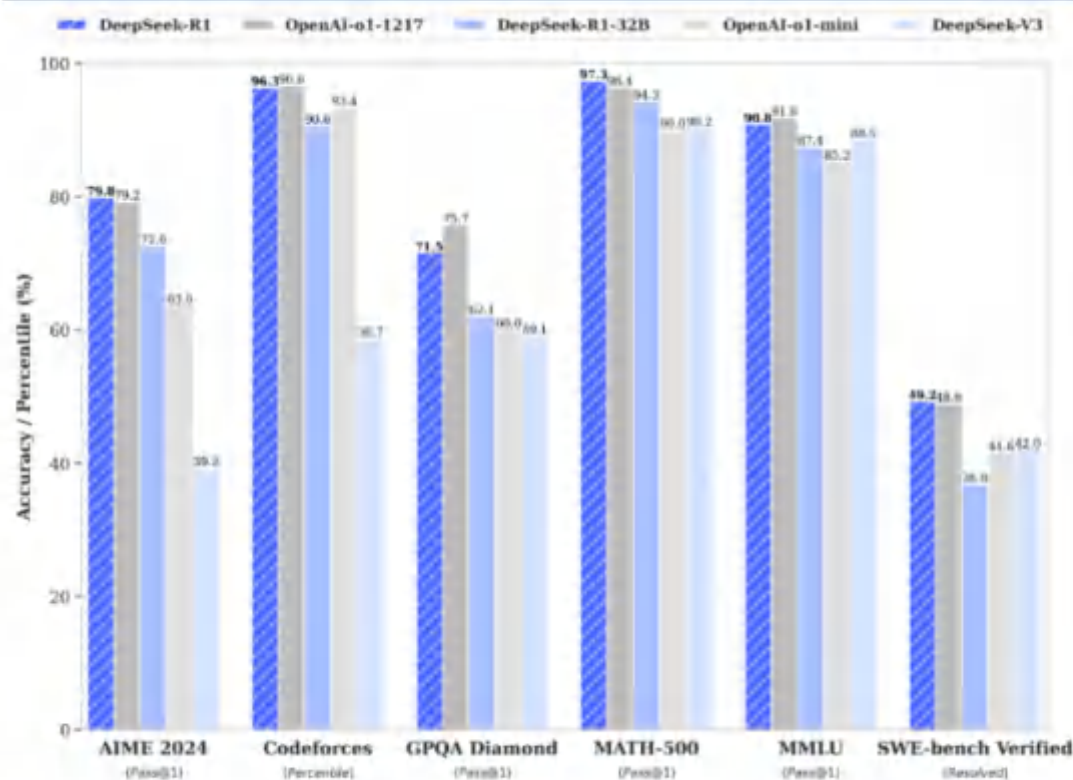


据介绍，R1模型在技术上实现了重要突破——用纯深度学习方法让AI自发涌现出推理能力，在数学、代码、自然语言推理等任务上，性能比肩OpenAI的o1模型正式版，且训练成本仅为560万美元，远低于美国科技巨头的数亿美元乃至数十亿美元投入。



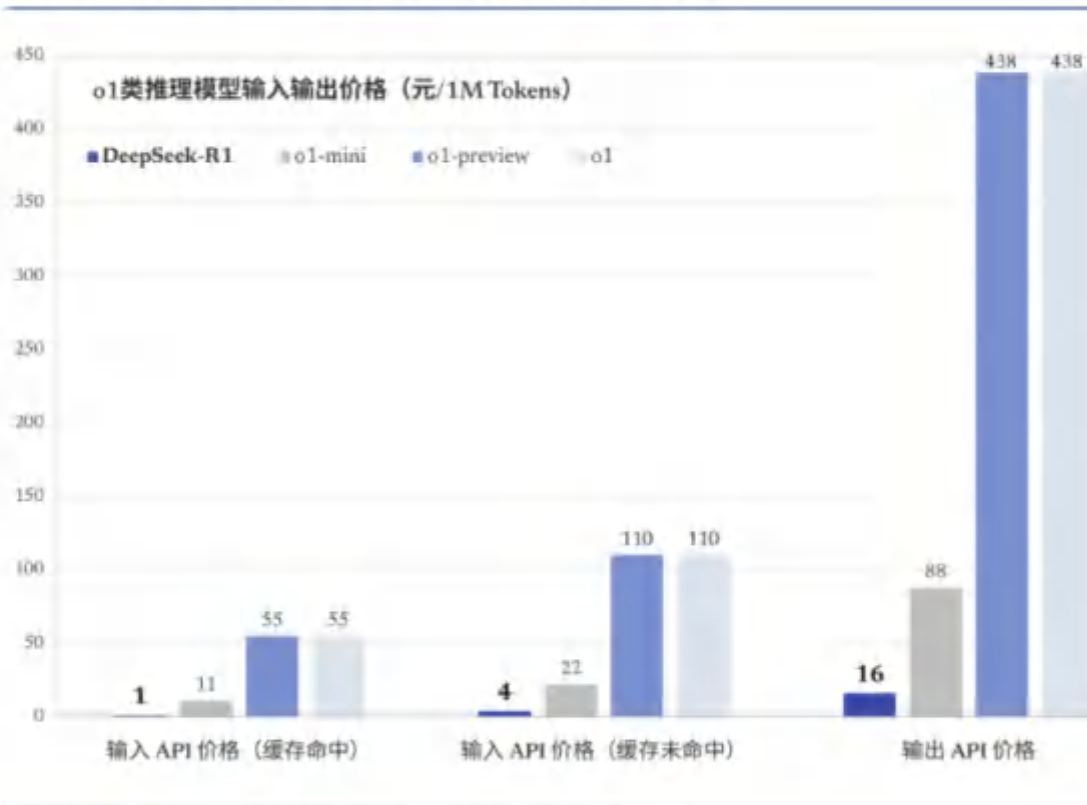
爆火原因

图2: DeepSeek 性能对齐 OpenAI-o1 正式版



资料来源: DeepSeek 官网, 中国银河证券研究院

图3: 推理成本低至每百万 Token 0.14 美元



资料来源: DeepSeek 官网, 中国银河证券研究院

- DeepSeek-R1的推理能力进入了第一梯队 (媲美OpenAI o1), 但训练和推理成本低、速度快、全部开源
- DeepSeek打破了硅谷传统的“堆算力、拼资本”的大模型发展路径

对AI行业的重大影响

01

打破垄断

DeepSeek-R1以低成本和开源特性打破以往头部企业巨头割据局面

02

价格下调

DeepSeek-R1的API定价仅为行业均价的1/10,推动了中小型企业低成本接入AI,对行业产生了积极影响

03

推动创新

DeepSeek-R1促使行业开始从“唯规模论”转向更加注重“性价比”和“高效能”方向

公司成立背景与发展历程

DeepSeek, 全称杭州深度求索人工智能基础技术研究有限公司, 成立于2023年7月17日, 是一家创新型科技企业, 专注于人工智能基础技术的研究与开发

大语言模型(LLM)的创新应用

DeepSeek专注于开发先进的大语言模型(LLM)和相关技术, 旨在通过这些技术推动人工智能在多个领域的应用和创新

投资者背景与市场定位

作为由知名私募巨头幻方量化孕育而生的公司, DeepSeek获得了强大的资金支持和行业影响力, 幻方量化与九坤投资、明汭投资、灵均投资并称量化私募领域的“四大天王”, 管理资金规模均超过600亿元。这为DeepSeek提供了清晰的市场定位和投资者背景

里程碑

2025年1月20日推出DeepSeek-R1推理模型

DeepSeek产品信息官方渠道

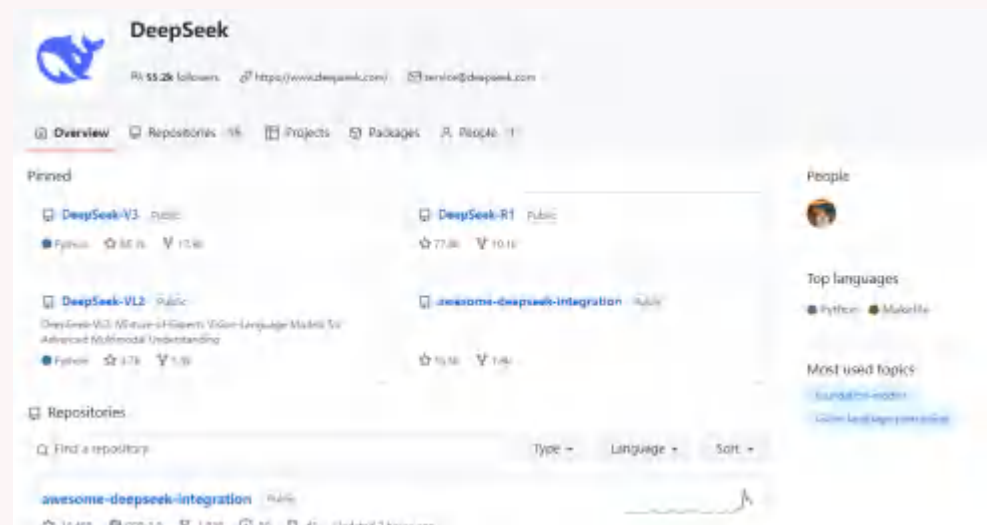


新闻: <https://api-docs.deepseek.com/zh-cn/news/news250120>

官网: www.deepseek.com

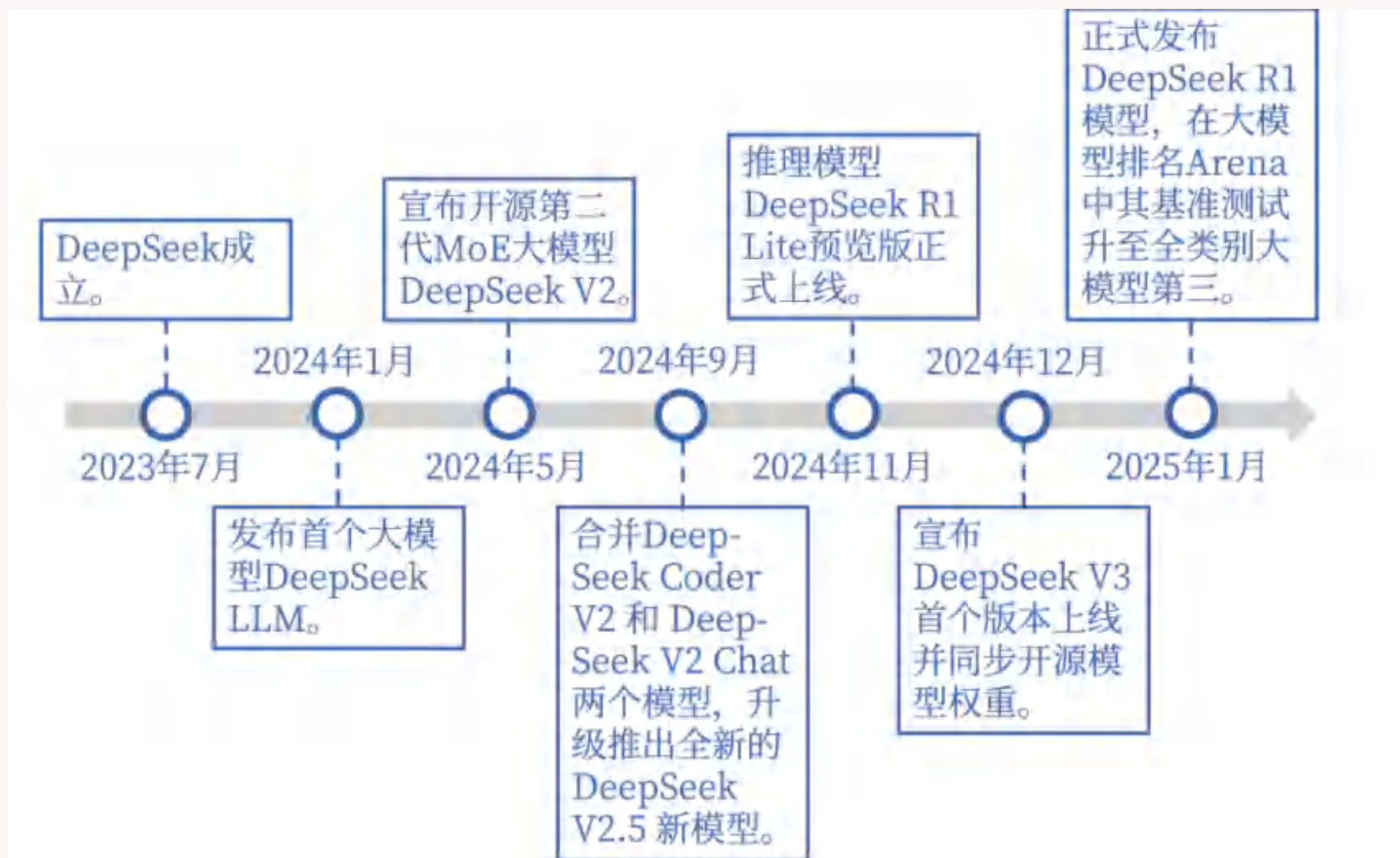


对话: chat.deepseek.com



Github: <https://github.com/deepseek-ai/>

DeepSeek模型系列



DeepSeek最新的生成模型和推理模型版本对比



比较方面	生成模型 (V3)	推理模型 (R1)
设计初衷	想要在各种自然语言处理的任务中都能表现好, 更通用	重点是为了搞定复杂的推理情况, 比如深度的逻辑分析和解决问题
性能展现	在数学题、多语言任务还有编码任务里表现不错, 像Cmath能得90.7分, Human Eval编码任务通过率是65.2%	在需要逻辑思考的测试里很棒, 比如DROP任务F1分数能达到92.2%, AIME 2024的通过率是79.8%
应用的范围	适合大规模的自然语言处理工作, 像对话式AI、多语言翻译还有内容生成等等, 能给企业提供高效的AI方案, 满足好多领域的需求	适合学术研究、解决问题的应用和决策支持系统等需要深度推理的任务, 也能拿来当教育工具, 帮学生锻炼逻辑思维

DeepSeek-R1 工作原理

思维链
(Chain of Thought)
让模型进行慢思考

强化学习
(Reinforcement Learning)
让模型自我探索和训练



蒸馏
(Distillation)
在不损失能力的情况下缩小模型

- DeepSeek R1论文: https://github.com/deepseek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf
- DeepSeek R1论文图解: <https://zhuanlan.zhihu.com/p/20844750193>

DeepSeek-R1 版本

模型名称	
DeepSeek-R1-671B	满血版, 能力最强
DeepSeek-R1-Distill-Qwen-7B	蒸馏版, 能力稍弱 --实际上是增加了推理能力的Qwen或Llama模型
DeepSeek-R1-Distill-Llama-8B	
DeepSeek-R1-Distill-Qwen-14B	
DeepSeek-R1-Distill-Qwen-32B	
DeepSeek-R1-Distill-Llama-70B	

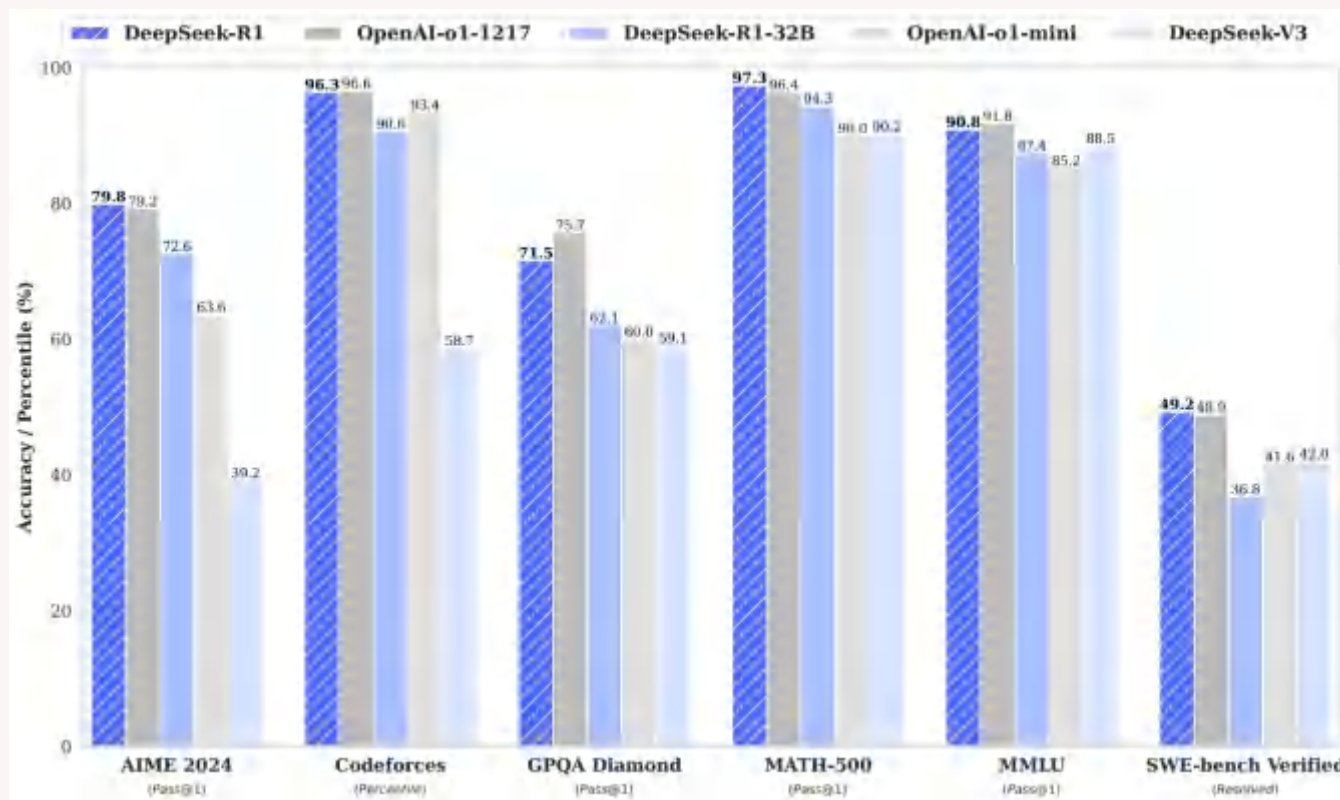
部署DeepSeek-R1 满血版的算力要求和性能

模型	算力型号	每套卡数	每套算力 (FP16/TFLOPS)	每套显存 (GB)	模型精度	含算力部署总价 (万/套/月)
 DeepSeek R1 671B	NV H800	16	31664	1280	FP8	18
	NV H200	8	15832	1128	FP8	11
	NV H20	16	2368	1536	FP8	8
	昇腾 910B	32	10016	2048	BF16	12
 DeepSeek V3 671B	NV H800	16	31664	1280	FP8	18
	NV H200	8	15832	1128	FP8	11
	NV H20	16	2368	1536	FP8	8
	昇腾 910B	32	10016	2048	BF16	12

DeepSeek-R1的优势

理科能力强，且准确率高

- 数学推理
- 代码生成
- 复杂任务处理



DeepSeek-R1 的局限

●通用能力

- R1 的通用能力（尤其是生成能力）低于DeepSeek-V3
- R1的幻觉仍旧比较明显（可能源于R1的中文语言表达能力更强）

●语言混杂

- R1 在处理非中英文问题时，偶尔会出现语言混杂现象
- 这个现象在R1 Zero版本中更加明显

●提示词工程

- 使用 few-shot 提示可能会降低R1性能
- 使用过多的过程指导指令可能会降低R1的推理能力

The results are shown in Table 1 below.

	DeepSeek R1	DeepSeek V3
Vectara's HHEM 2.1	14.3%	3.9%
Google's FACTS w/ GPT-4o & Claude-3.5-Sonnet	4.37%	2.99%
Google's FACTS w/ GPT-4o & Gemini-1.5-Pro	3.08%	1.99%
Google's FACTS w/ Claude-3.5-Sonnet & Gemini-1.5-Pro	3.89%	2.69%

Table 1: Hallucination rates of DeepSeek R1 and V3 by various hallucination judgment approaches. Lower hallucination rates are better.

Thus our surprise: consistently across all judgment approaches, Deepseek-R1 is shown to be hallucinating at significantly higher rates than Deepseek-V3.

DeepSeek-R1 应用场景



- 推理密集型任务

- 编程任务中的代码生成、算法设计，媲美Claude 3.5 Sonet
- 数学问题求解、科学推理和逻辑分析等需要复杂推理的场景。

- 教育与知识应用

- 可用于解决教育领域的问题，支持知识理解与解答。
- 可用于科研任务的实验设计、数据分析和论文撰写。

- 文档分析与长上下文理解

- 适合处理需要深入文档分析和理解长上下文的任务，例如复杂信息提取与整合。

- 开放领域问答与写作

- 在内容生成、问题回答以及创造性写作中具有广泛应用，例如生成高质量文本或进行内容编辑。

如何使用DeepSeek R1

- DeepSeek官方网站和官方app

- 腾讯系

- 腾讯元宝
- 微信AI搜索
- 腾讯ima个人知识库

- AI搜索

- 秘塔搜索
- 纳米AI搜索
- 知乎直答

- 其它接入DeepSeek-R1的产品

- Molly R1 (向量智能)
- 问小白



接入DeepSeek R1第三方服务的厂家



- 互联网大厂：
 - 腾讯：腾讯元宝、ima、微信、腾讯云
 - 百度：百度搜索、文小言
 - 字节跳动：豆包、扣子、飞书
 - 阿里：钉钉
 -
- 互联网小厂：科大讯飞、知乎、秘塔、纳米AI搜索、.....
- AI初创公司：零一万物、阶跃星辰、minmax、.....
- AI应用公司（教育类）：学而思、北大青鸟、网易有道、猿辅导、作业帮、.....
- 手机厂商：华为、荣耀、魅族、.....
- 其它云服务平台：三大运营商、云服务商（硅基流动等）



PART 02 ▶

AIGC的概念和应用

本部分着重介绍人工智能生成内容（AIGC）的定义、应用范围及其在各行业的实际影响：

首先梳理AIGC相关的术语，包括AI、AGI、生成式AI和决策式AI等，明确了其在人工智能领域的定位。随后，详细列举AIGC在文本、图像、音频和视频生成方面的多样化应用，并探讨其在电商、新闻传媒、影视、游戏、教育和金融等行业的具体实践，展示AIGC在提升效率、降低成本和增强创新方面的核心价值。

此外，还分析AIGC带来的挑战，如数据隐私、伦理问题、生成质量控制以及对就业结构的影响。通过深入探讨AIGC的应用现状和未来趋势，本部分旨在帮助读者全面理解AIGC的潜力和影响，为应对技术变革提供参考。

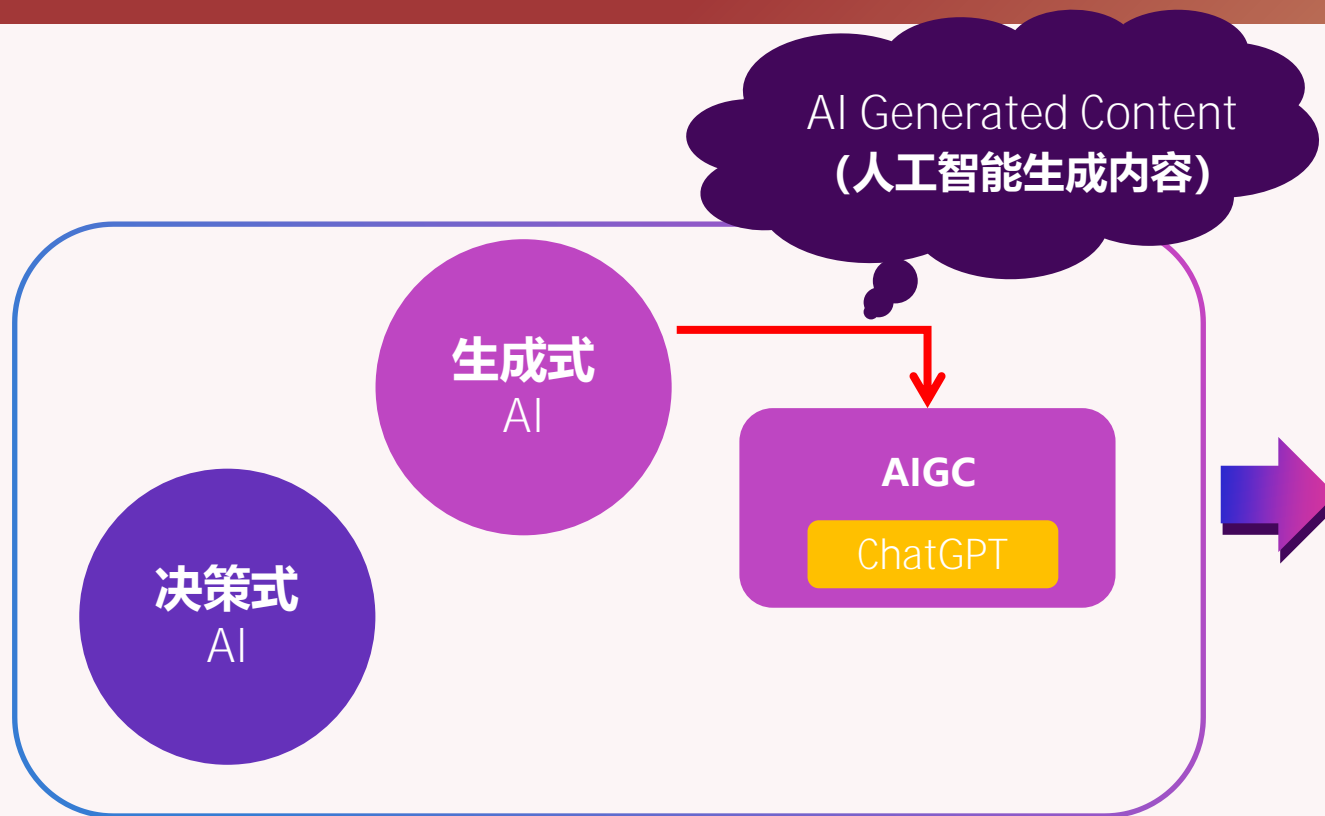
本部分的内容参考了《人工智能通识课（微课版）》和相关的视频（B站的思睿观通）

初识 AIGC

眼花缭乱的名词



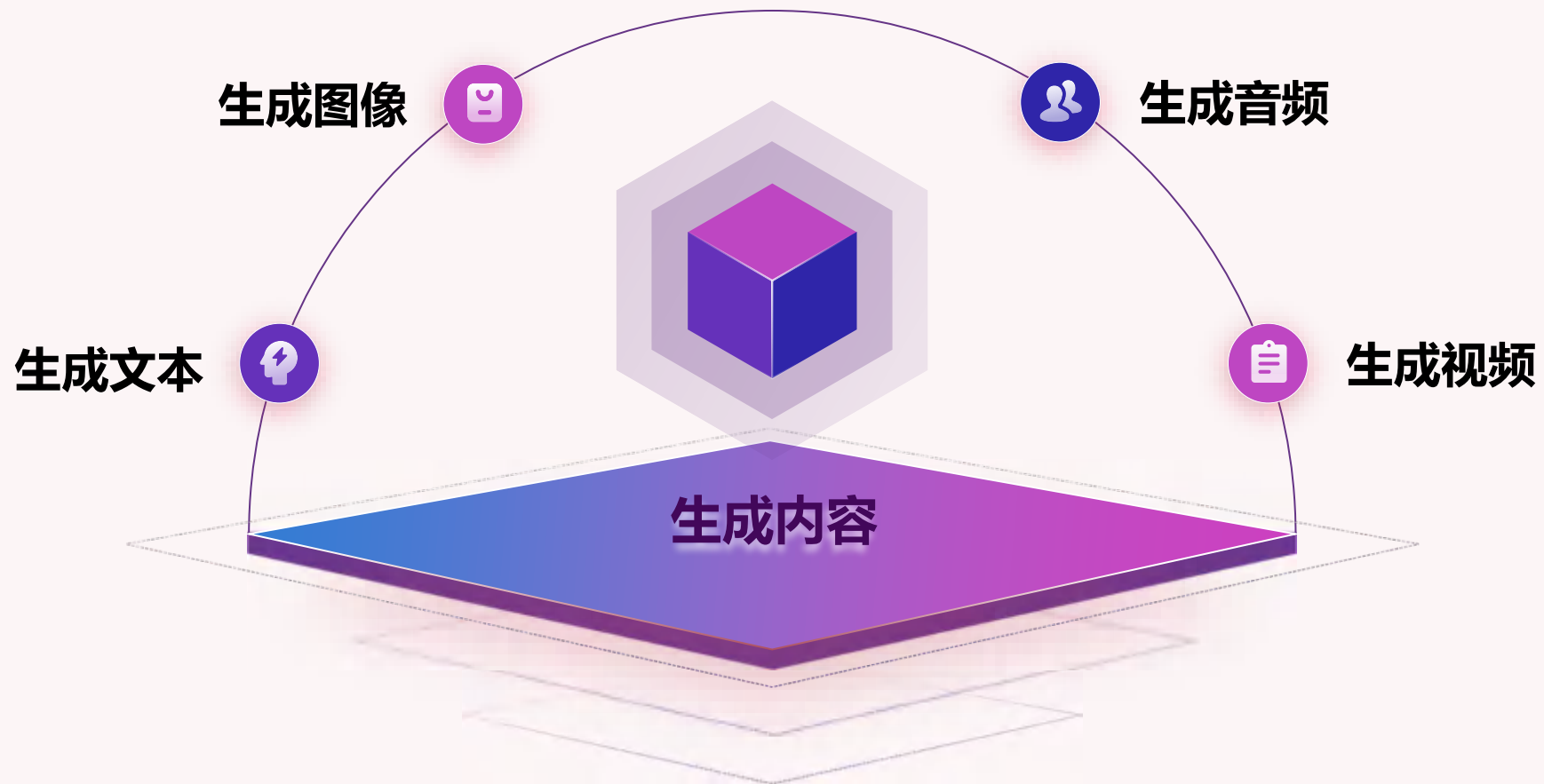
眼花缭乱的名词



AI(人工智能)

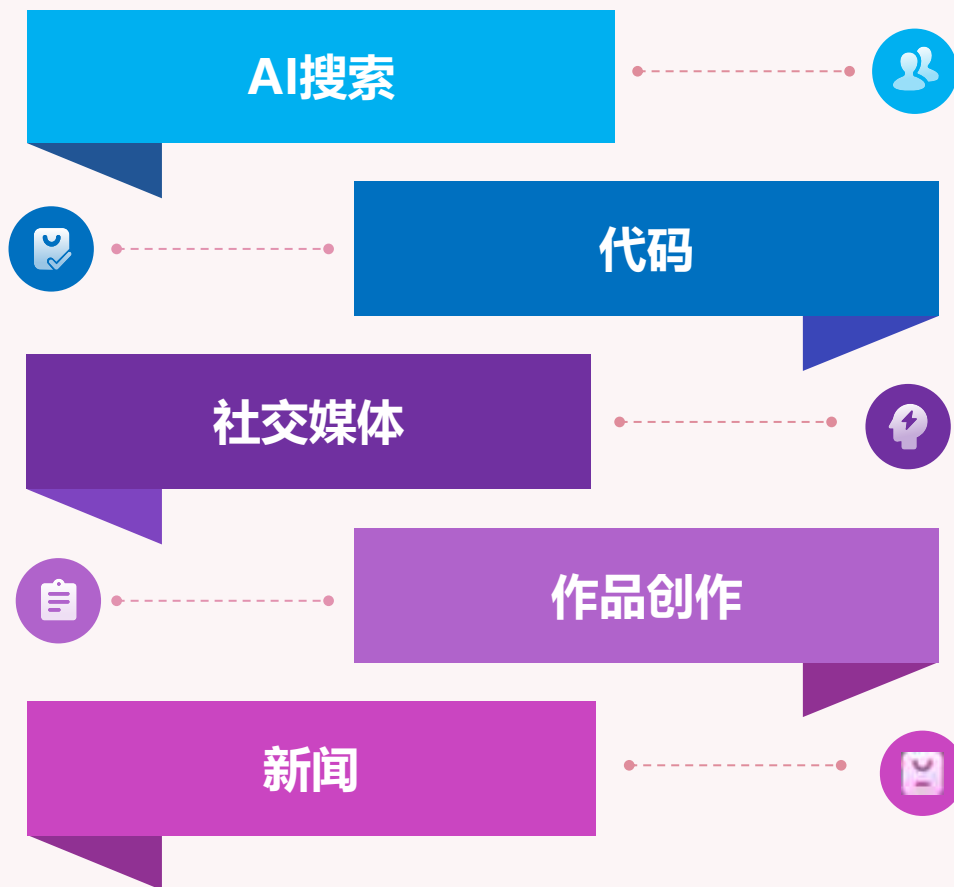
AGI(通用人工智能)
(artificial general intelligence)

AIGC能干什么



- Cursor
- Windsurf
- MarsCode
- GitHub Copilot
- 阿里云通义灵码

- 微软“小冰”出版诗集
- AI撰写剧本电影《阳春》
- 百万字小说《天命使徒》
- 清华大学的《机忆之地》



- Perplexity
- 知乎直答
- 秘塔AI搜索引擎
- 纳米AI搜索
- 微信搜索 (DeepSeek R1、元宝)
- 小红书的笔记助手
- 短视频平台的内容创意
- 自动回复机器人
- 企业微信的智能回复
- 英国广播公司(BBC)的Juicer
- 彭博社的Cyborg
- 新华社的快笔小新
- 人民日报的创作大脑
- 第一财经的DT稿王

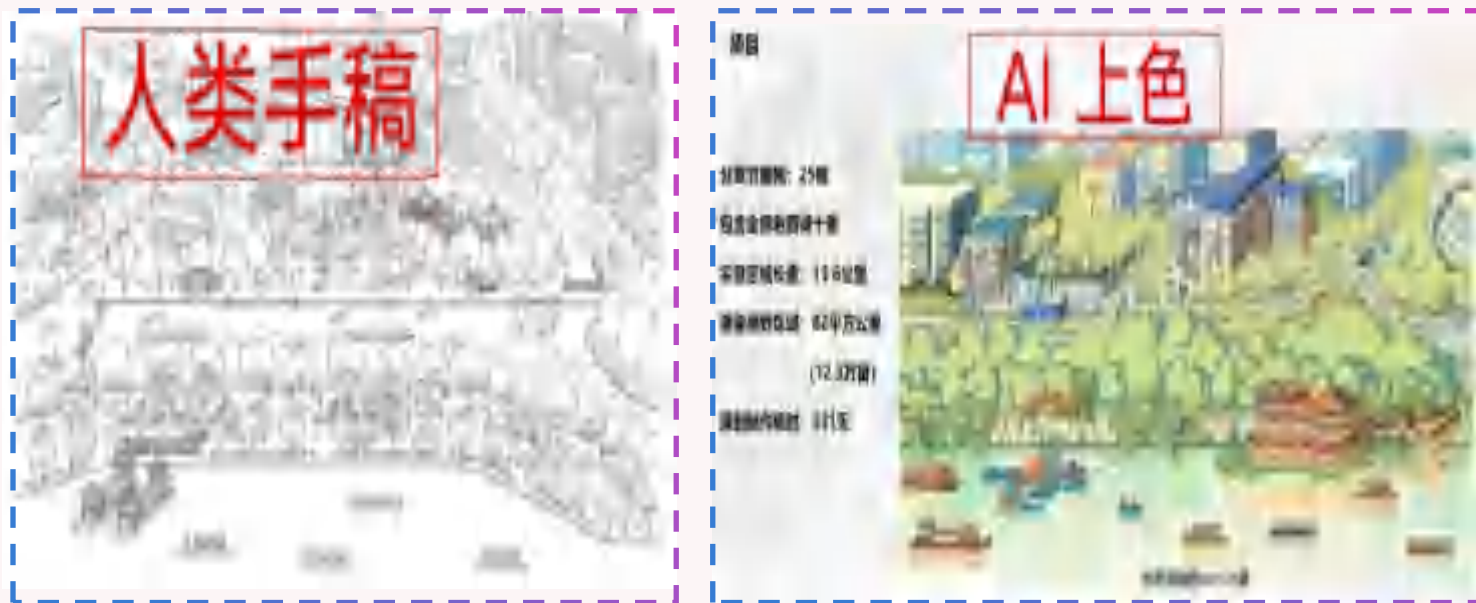


太空歌剧院

2022年8月，获得比赛
“数字艺术/数字修饰照片”
一等奖

AI绘画



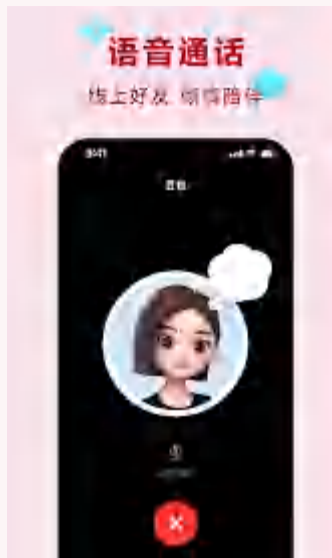


全景图是一副百米画卷，再现了当下西湖边的景观特色、生活场景，还有一些已经消失的景观，比如曾经闻名遐迩的大佛寺。

图中包含5000个建筑，作者白小苏说如果这些建筑都要自己亲手画，一天画10个都要500天。最后他借助于AI技术，在一年内完成了作品



新西湖繁胜全景图



语音对话



创作音乐



语音克隆



千秋诗颂



视频特效



视频添加图片

AIGC 的行业应用

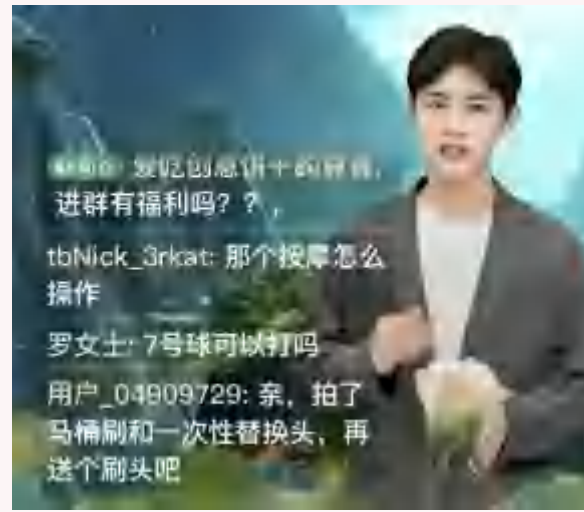




商品3D模型
改善购物体验

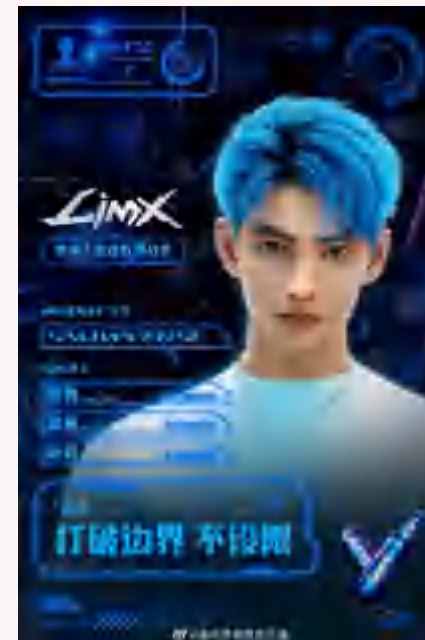
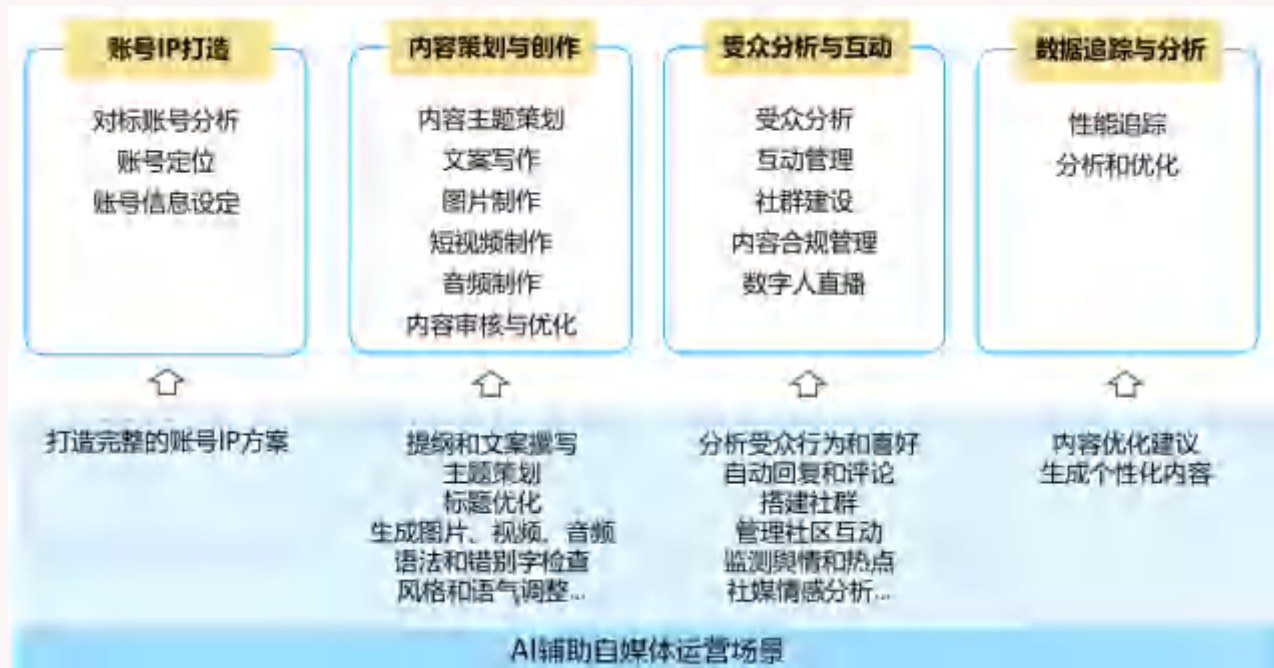


AI模特
产品拍摄降本增效



虚拟主播
提升直播带货效率

内容广告



**全流程参与
提高效率**

**雪花啤酒虚拟偶像
提升品牌宣传效果**

AIGC对传媒行业影响包括采编环节、传播环节，通过语音转写、智能写作、智能剪辑等方式提高采编环节的生产效率，在传播环节通过打造AI主播实现智能、高效播报。



AIGC对传媒行业参与主体的影响



智能图像修复



文本生成场景



前期创作

影视剧本文稿创作

通过对海量剧本数据进行分析归纳，并按照预设风格快速生产剧本，创作者再进行筛选和二次加工，激发创作者的灵感，缩短创作周期。

影视剧本创作平台



中期拍摄

拓展角色范畴

通过AIGC合成人脸、声音等内容，实现“数字复活”已故演员，替换“劣迹艺人”、高难度动作合成等

拓展场景空间

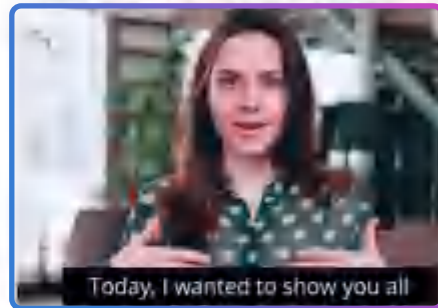
AIGC合成虚拟物理场景，将无法实拍或成本过高的场景生成出来，大大拓宽了影视作品想象力的边界，给观众带来更优质的视觉效果和听觉体验

后期制作

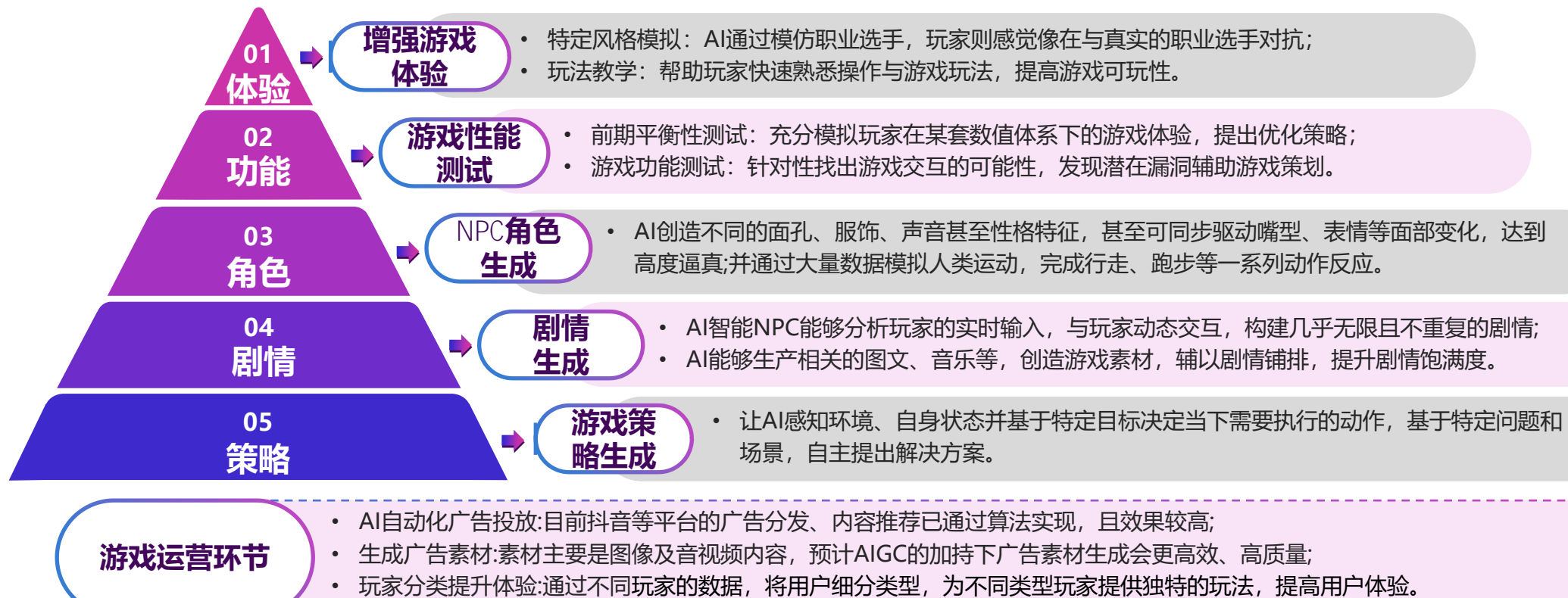
赋能影视剪辑

修复、还原影像，提升影像资料的清晰度；实现影视预告片自动生成；实现将影视内容维度转制，从2D向3D自动转制

实时字幕



AIGC 游戏研发环节作用



教育行业



其它

- 自动批改作业和考试
- 虚拟实验室
- 生成学习资料
- 翻译文献资料

金融行业

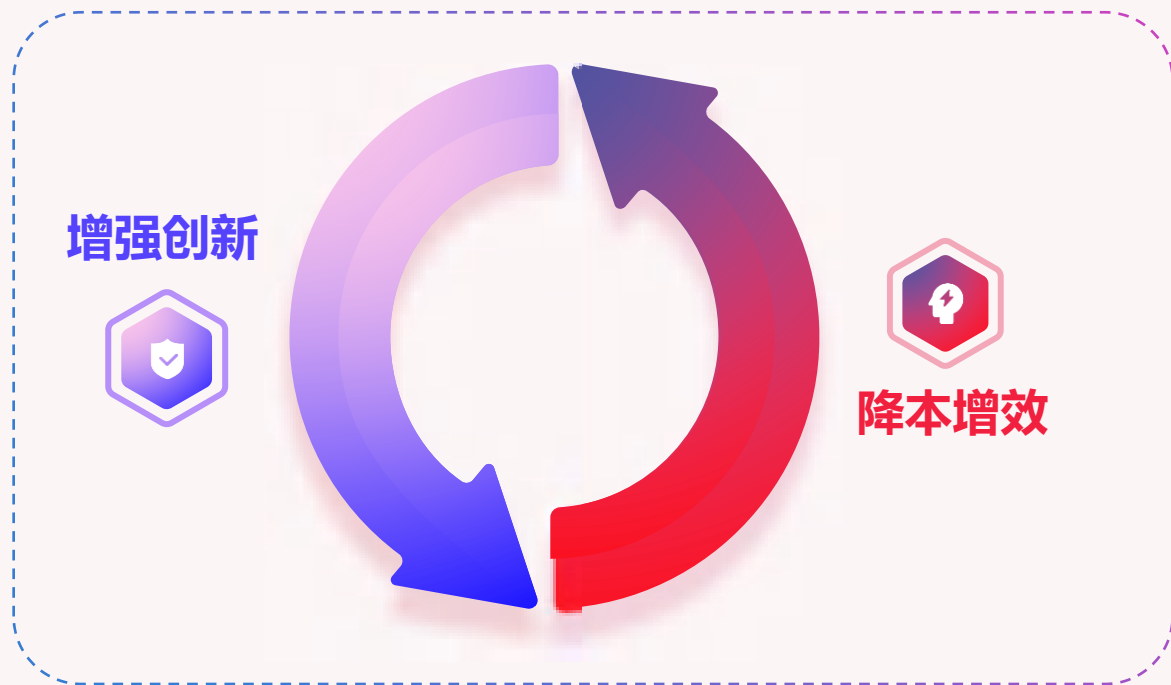


其它

- 风险管理
- 自动处理和分
析金融文档
- 合规管理

AIGC 的未来

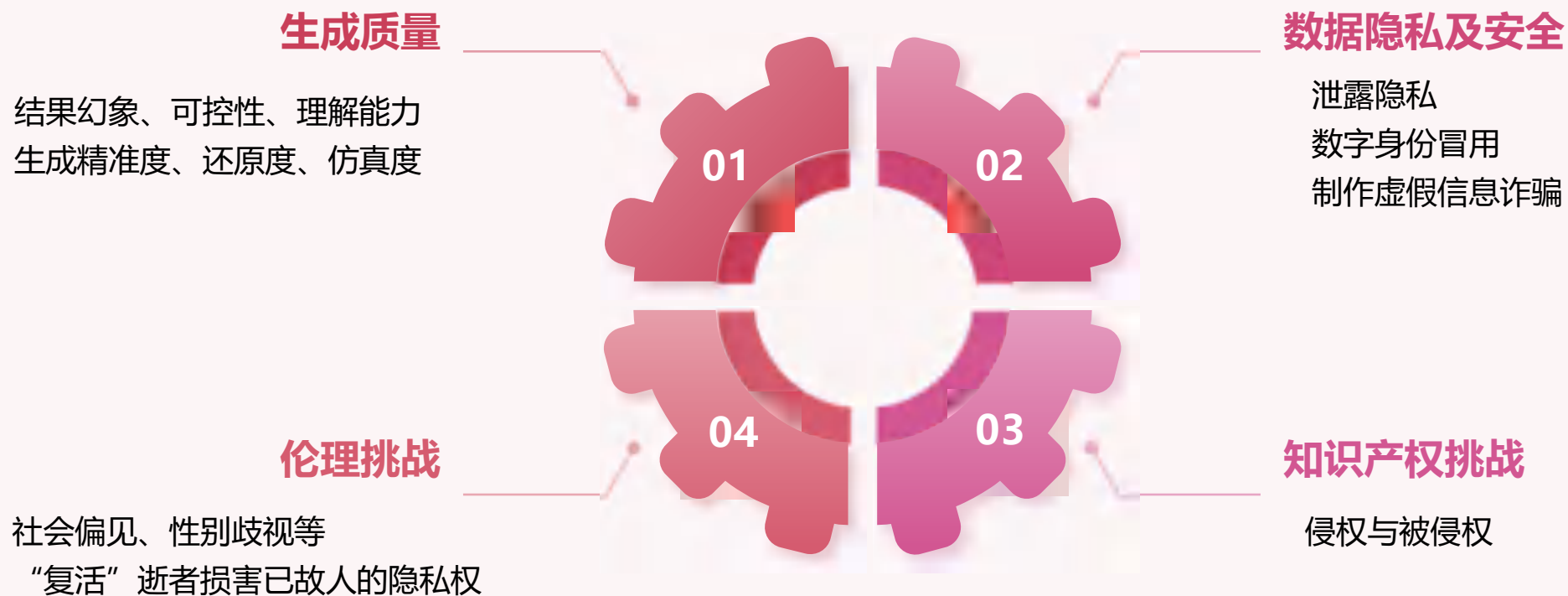
AIGC的核心价值



企业赋能



个人赋能



01

技术进步

AIGC的生成质量和准确性将得到提升

02

更广泛的应用

在更多领域得到应用，
如保险、医疗和法律
服务等

03

人机协作

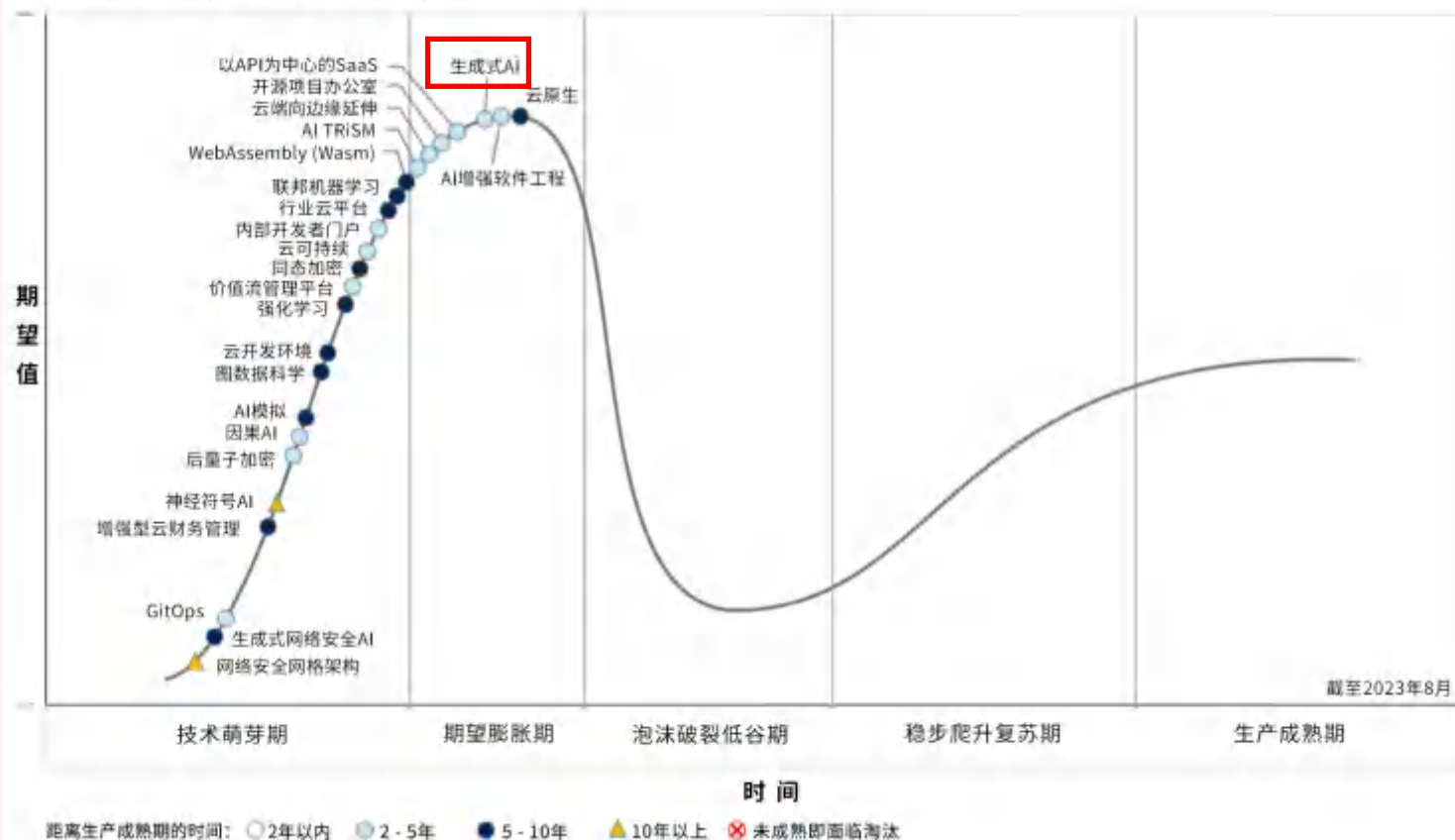
更多地与人类专家协作，
以提供更高质量
的内容和服务

04

监管框架的发展

预计将出现更多关于
内容生成和使用的监
管框架

2023年新兴技术成熟度曲线



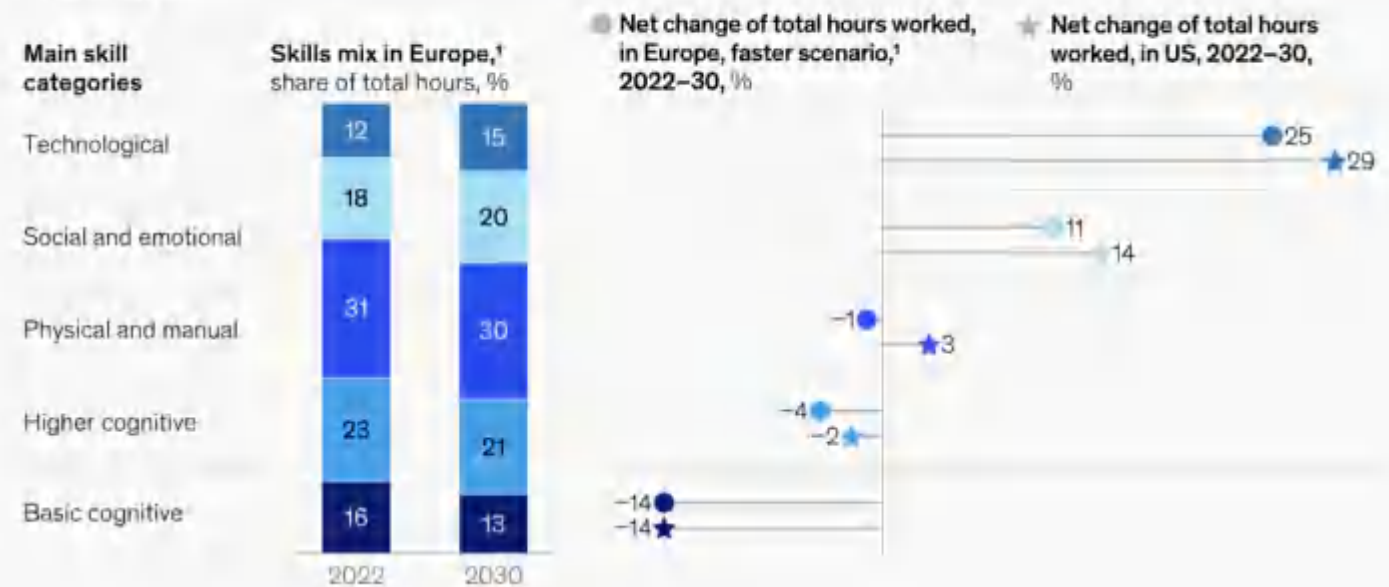
生成式人工智能 (AI) 目前处于**期望膨胀期**，预计将在2~5年内产生巨大效益

到 2026 年，Gartner预测超过**80%**的企业将使用生成式AI的API或模型，或在生产环境中部署支持生成式AI的应用，而在2023年初这一比例不到5%。

数据来源：Gartner (2023年8月)

技能类型需求变化

Demand for technological and social and emotional skills could increase in Europe.

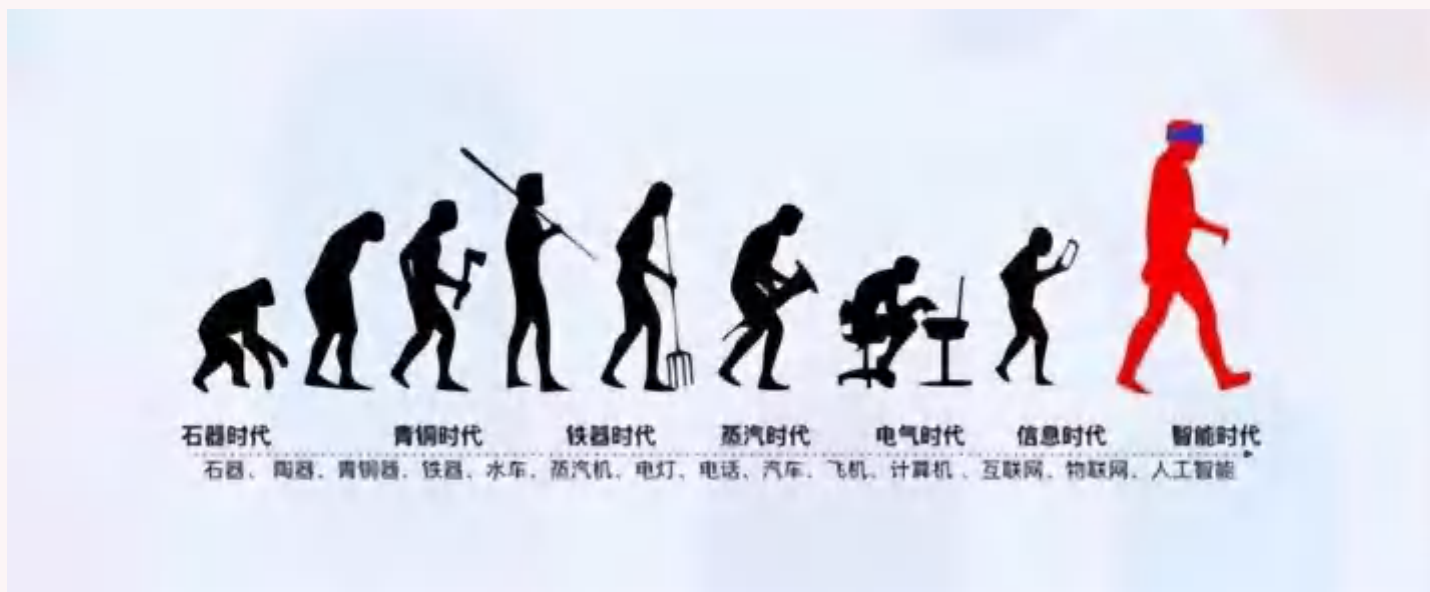


技能类型需求变化

- **技术技能**大幅增长
- 炙手可热的“新宠”技能是**社会和情感技能**
- 体力和手动技能的需求预计将保持大致稳定
- 认知技能(文字和信息处理、编程、科研、工程等)的需求预计将减少14%

2024年5月24日

麦肯锡报告：《工作的新未来：在欧洲及其他地区部署人工智能和提升技能的竞赛》



- 学习AIGC基础知识
- 积极使用AIGC工具
- 关注AIGC在各行业的应用案例
- 跟踪最新发展趋势



PART 03 ▶

AIGC的能力揭秘

简介



本部分深入剖析AIGC背后的技术原理，涵盖文本生成和图像生成两大核心领域。通过对比文本生成和图像生成的不同路径，大家能够全面理解AIGC在不同模态下的工作原理，以及如何利用这些技术实现高效的内容创作和应用。

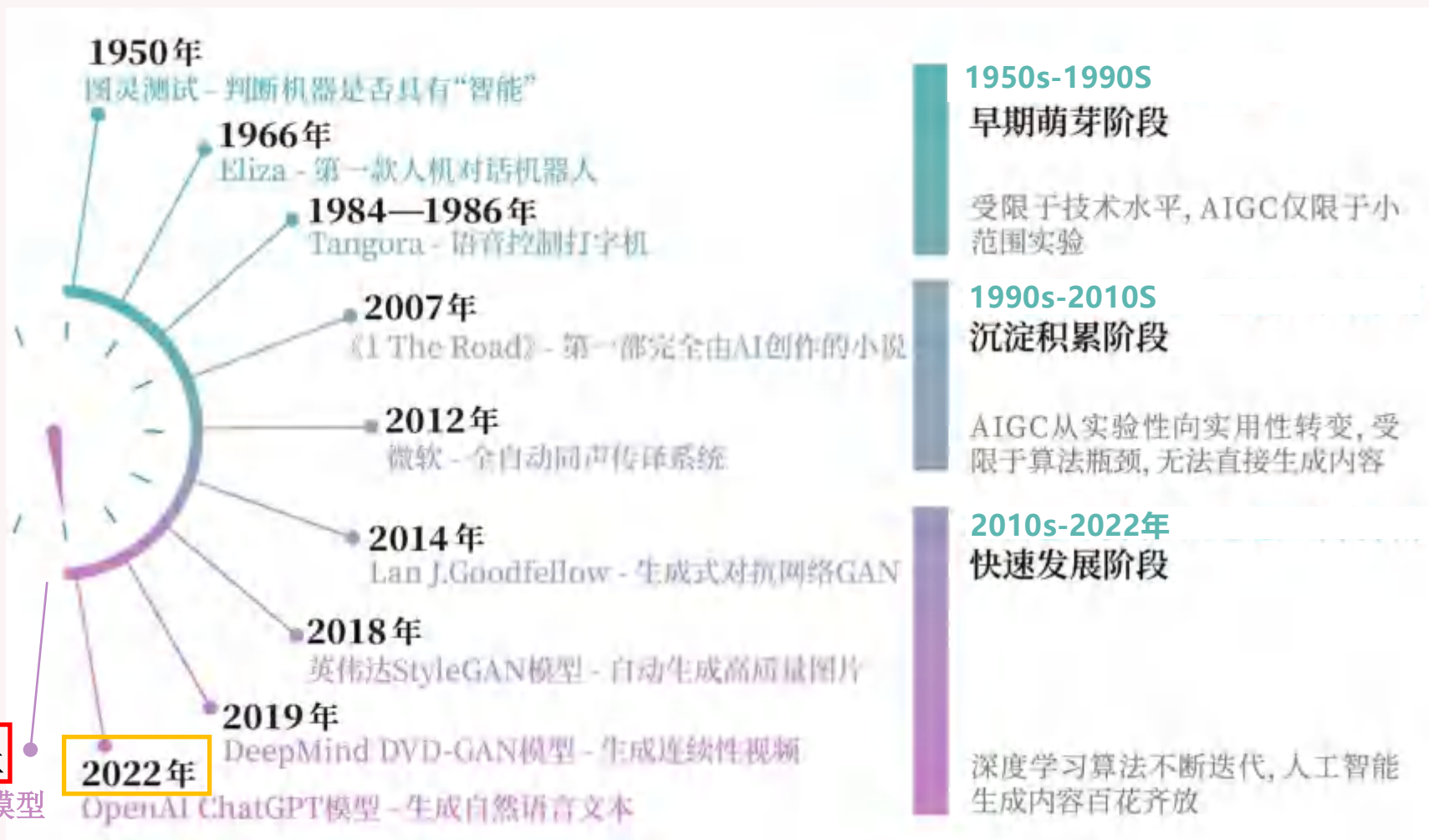
在文本生成方面，以OpenAI的GPT-4o为例，详细介绍了通用大语言模型（LLM）的工作原理，包括其基于Transformer架构的生成机制、上下文编码、自注意力机制以及预训练和微调过程。探讨了GPT-4o在多轮对话、语言转换、意图理解、文本生成和推理能力等方面的优势，同时也指出其在知识局限、上下文窗口限制和生成幻觉等方面的不足。

在图像生成方面，以Stable Diffusion模型为例，解释了其核心组件（文本编码器、图像信息生成器和图像解码器）的工作流程，揭示了文生图和图生图技术的优势与局限。

本部分的内容参考了《人工智能通识课（微课版）》和相关的视频（B站的思睿观通）。

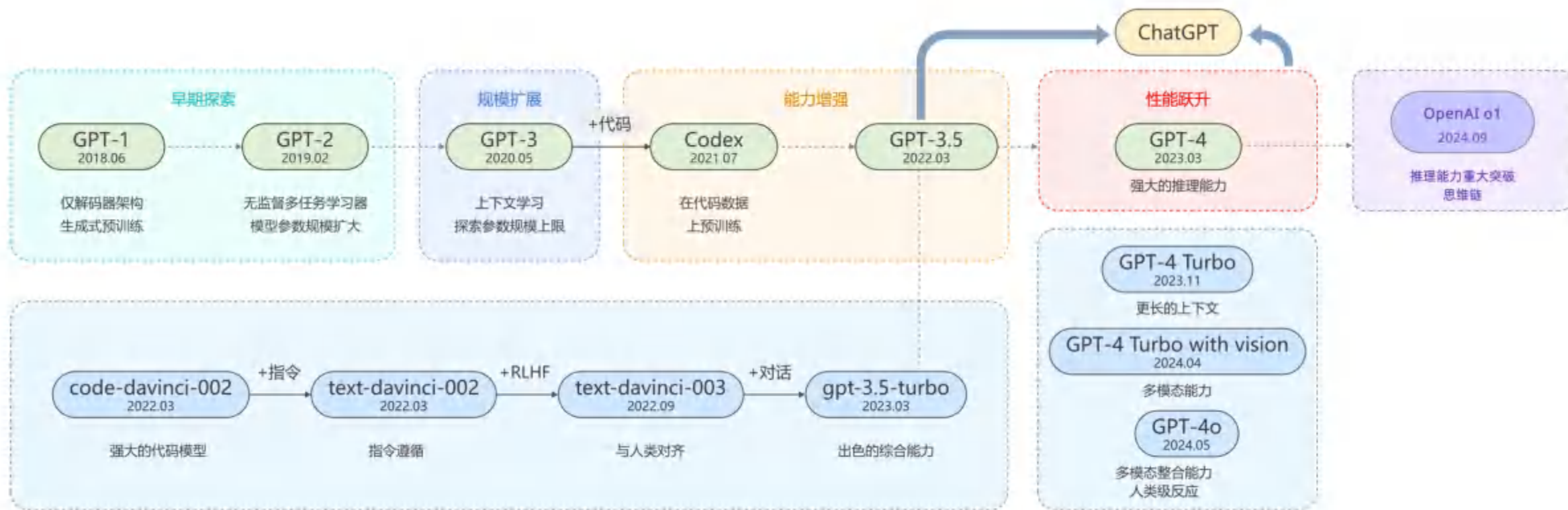
文本生成的奥秘

AIGC与GPT (通用大模型的代表)



OpenAI GPT的学霸养成记

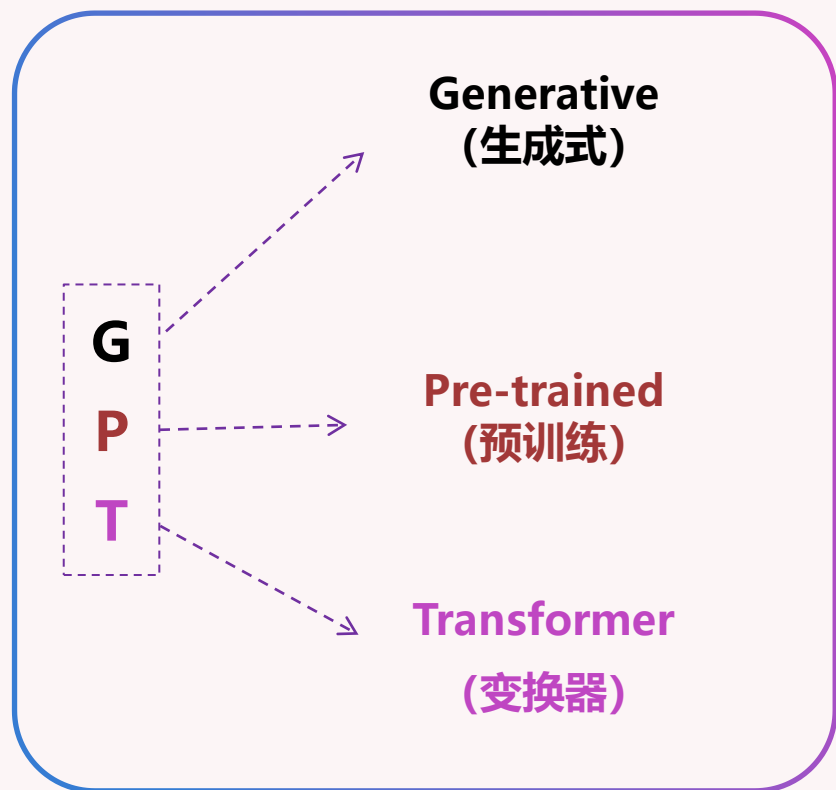
GPT:生成式预训练变换模型(Generative Pre-trained Transformer)



3大核心能力



工作原理-1



LLM: Large Language Model

1. 收到提示词

示例: “今天天气不错, 我决定”

2. 将输入拆分为token

[“今”, “天”, “天”, “气”, “不”, “错”, “,”, “我”, “决”, “定”]

3. 采用Transformer架构处理token

- 理解token之间的关系
- 识别提示词的整体含义

4. 基于上下文预测下一个token

- 为可能的单词分配概率分数
- 示例: { “去” :0.7, “停”:0.2, “站”:0.1 }

5. 根据概率分数选择标记

示例: “去”

概率预测+文字接龙

自回归: 重复步骤4和步骤5直到形成完整的句子

示例: 今天天气不错, 我决定去公园

工作原理-2

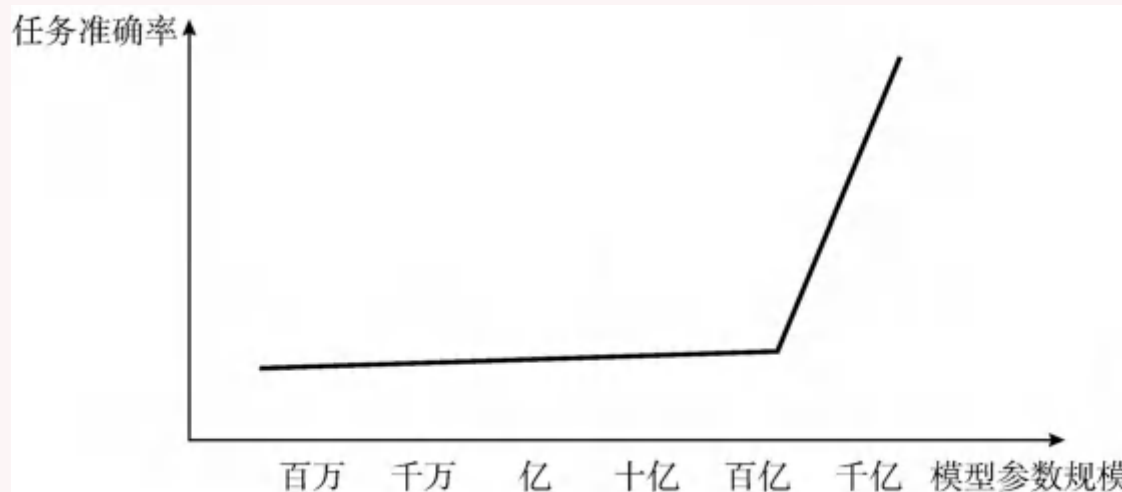
大模型工作过程



工作原理-3

数据来源	说明
维基百科	在线百科,严谨
图书	经典为主的古登堡计划和自助出版平台Smashwords等
杂志期刊	论文: ArXiv等
链接	WebText,Reddit
Common Crawl	开源项目,爬取互联网所有数据
GitHub	程序员聚集地
合计	<ul style="list-style-type: none">700 多GB,约有19万套四大名著的阅读量5 000亿左右的token数量。(13 万亿token:gpt4)100个标记大约等于75个英语单词

训练语料 (gpt-3)



典型的新技能学习曲线: 规模到达临界点之后才会迅速增长

模型参数: 1.8万亿参数 (GPT-4)

GPT-4o

上下文窗口大小: 8192个token(标记)

生成式通用大语言模型的优势与劣势

优势

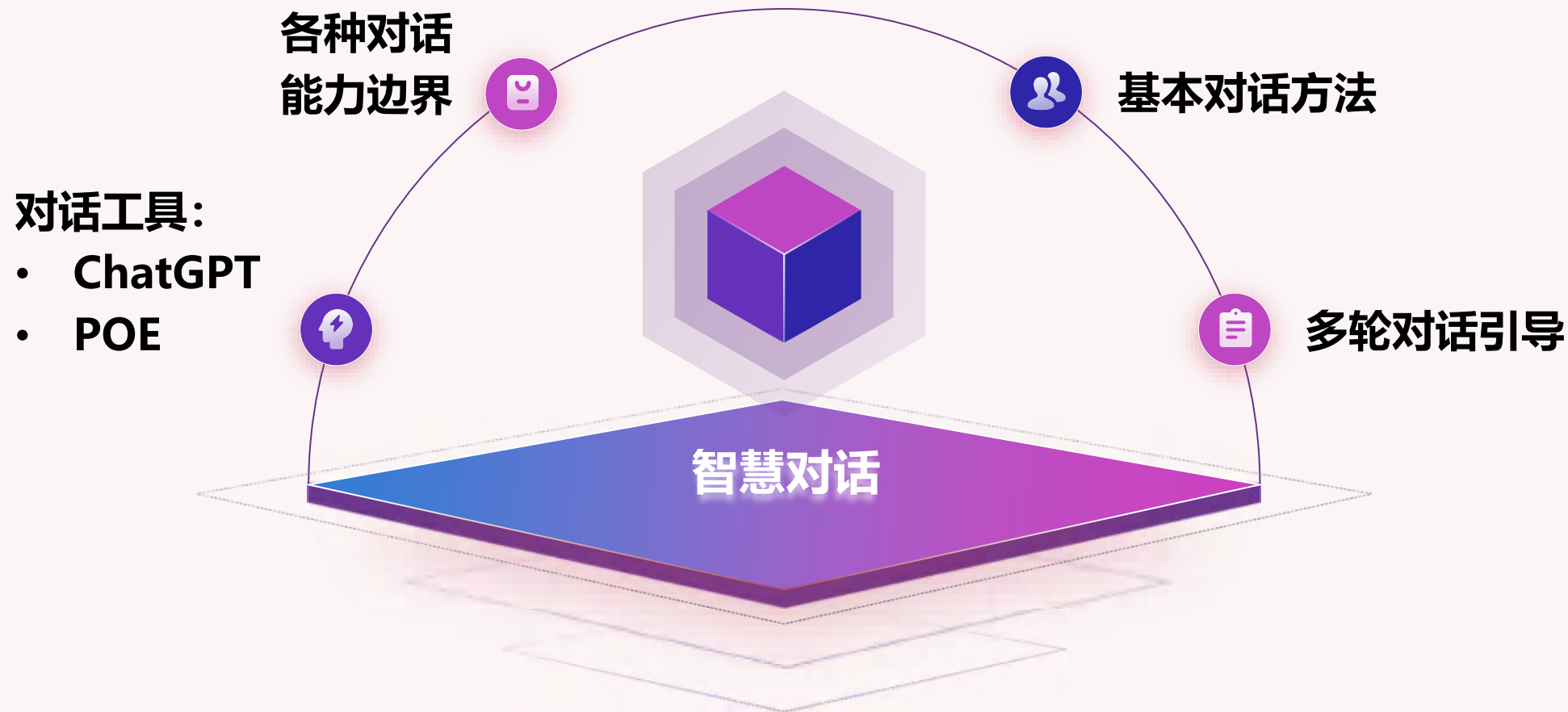
- 语言理解和生成能力
- 世界知识能力
- 一定的推理能力

VS

劣势

- 幻觉（生成错误答案）
- 知识库有限
- 上下文窗口限制

如何与GPT对话



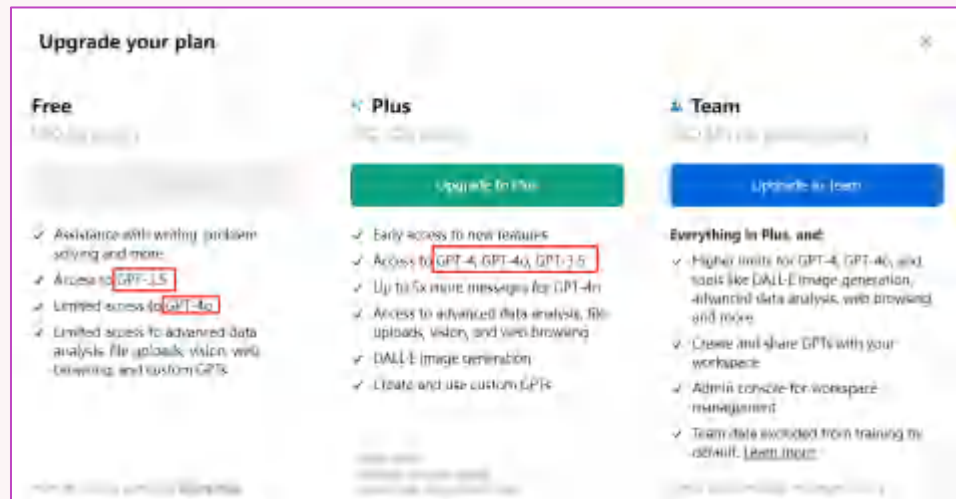
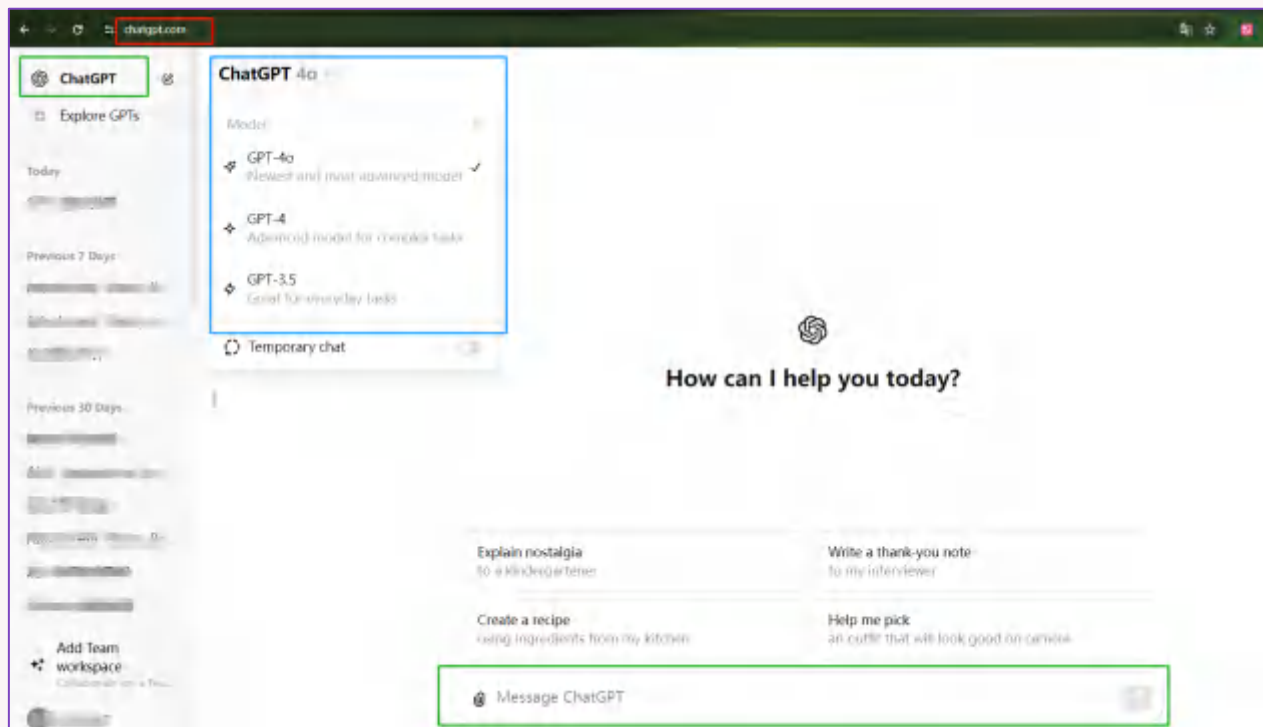
使用对话工具

ChatGPT: 聊天机器人应用

访问地址: <https://chatgpt.com/>

GPT: 大语言模型

GPT-3.5、GPT-4o



对话能力-多轮对话

原理

- 上下文编码
- 自注意力机制

限制

上下文窗口限制

- 问题+回答总数
- 包括字符、标点、空格

使用建议

- 分段对话
- 定期总结
- 使用关键词提醒
- 精简输入

怎么给刚满过十岁生日? → ChatGPT → 可
怎么给刚满过十岁生日? 可 → ChatGPT → 40
怎么给刚满过十岁生日? 可以 → ChatGPT → 通
怎么给刚满过十岁生日? 可以通 → ChatGPT → 可
怎么给刚满过十岁生日? 可以通以 → ChatGPT → 40
怎么给刚满过十岁生日? 可以通以以 → ChatGPT → 下
怎么给刚满过十岁生日? 可以通以以于 → ChatGPT → 可

- GPT-3.5: 4096, 大约3000 个英文单词或 4000 个汉字
 - GPT-4: 8192, 大约6000 个英文单词或 8000 个汉字
- (<https://platform.openai.com/tokenizer>)

不同语言

- 人类语言之间转换
- 机器语言之间转换
- 人类语言与机器语言之间转换

限制

- 人类语言生成能力不同
前5名：英语、西班牙语、法语、德语、中文
- 机器语言生成能力不同
前4名：python、Java、Typescript、Javascript

使用建议

- GPT:使用英文
- 中文优先选择国内产品
- 编程尽量选用前4种语言

对话能力--意图和情感分析能力

能力

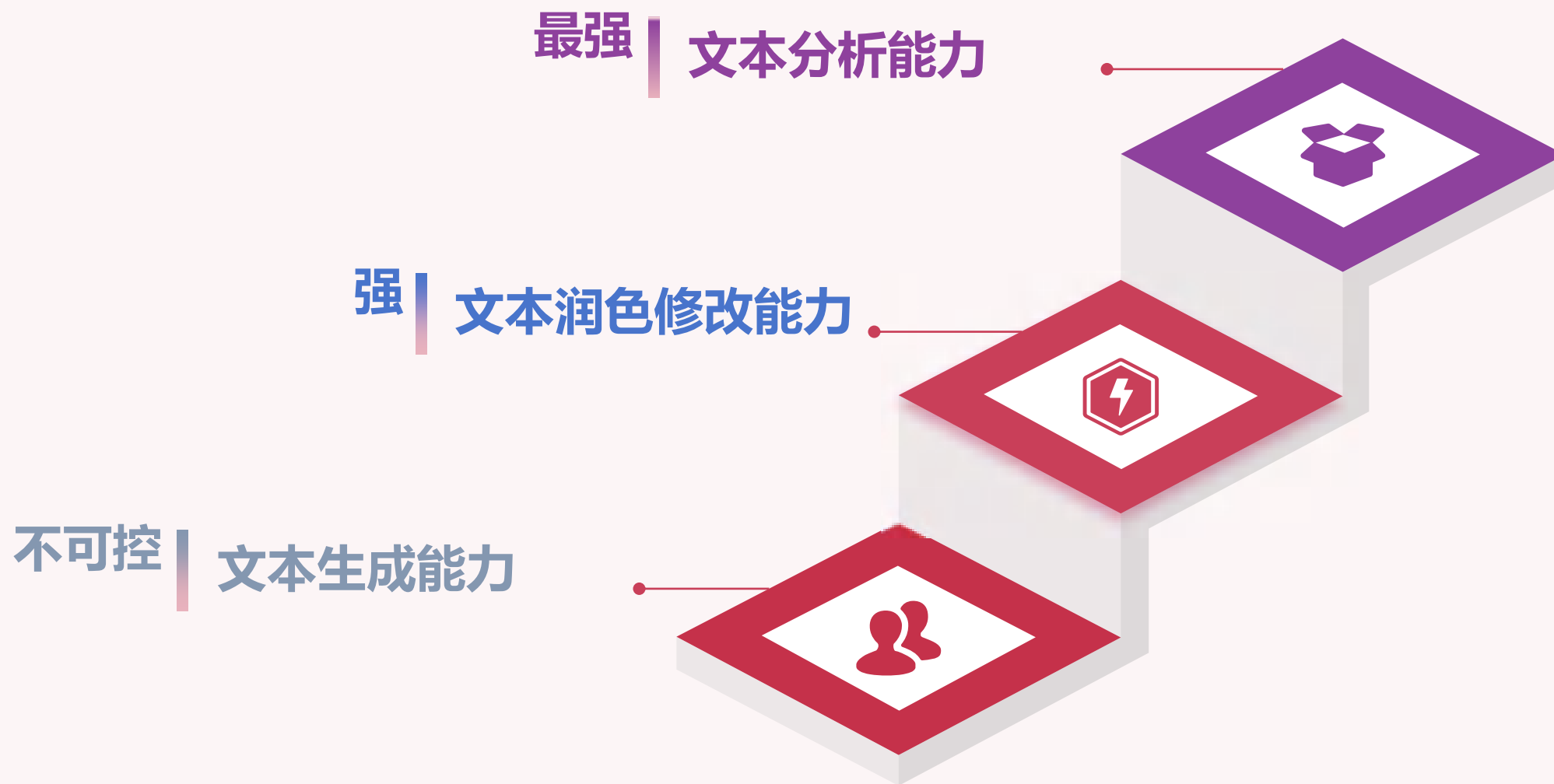
- 真实意图
- 情绪识别
- 判断行为特点
- 重要他人

限制

- 缺乏真正的情感
- 上下文理解有限
- 缺乏深层次推理

使用建议

- 提供清晰具体的信息
- 简短和连贯的对话
- 使用明确的情感表达





文本统计

- 统计字符数、单词数、句子数等基本统计信息



文本摘要

- 从长文档中提取主要信息，并生成一个简短的概述



文本分类

- 新闻分类、情感分析（正面、负面、中性）



文本总结

- 文本的整体内容进行简要概述，通常比文本摘要更为简短

限制

知识更新、长文本理解、数据偏见、数据质量、语境歧义、领域专业知识、新颖和未知类别、不确定性



文本校对

- 错别字识别与修改、语法错误检查与修正、标点符号调整、词汇选择优化、语句重组与调整、保持一致的写作风格



文本改写

- 口语转书面语、不同体裁风格、语义保持、文本简化、文本优化、保持一致性



文本扩写

- 增加细节、举例说明、提供背景信息、解释专业术语或概念、描述过程或步骤、增加文学修辞、提供多种观点等

限制

新词汇与表达、语境误判、高度主观性、文化敏感性和幽默、不断更新的语言规则、知识更新限制、长文本处理能力、用户意图识别



使用优先级

- 文本分析 > 文本润色 > 文本生成



分段处理长文本



明确和具体的指示



补充背景信息



注意多义词和模糊表达



对比能力

- 对比分析：事物、概念、人物、事件、文学作品、艺术作品等



列举能力

- 简单列举任务：组合/排列/序列，简单数学问题等



假设推理能力

- 多个领域能力不同：历史事件、文学作品解读等

限制

主观判断、上下文窗口、非专业算法、数据局限、缺乏常识和领域专长、模型理解能力、过度生成偏见、难以验证答案的准确性

使用建议

- 明确问题和期望
- 补充背景信息
- 结合领域知识

基本对话



学会提问

- 明确、具体、避免歧义
- 逻辑清晰
- 上下文信息
- 示范数据



验证输出

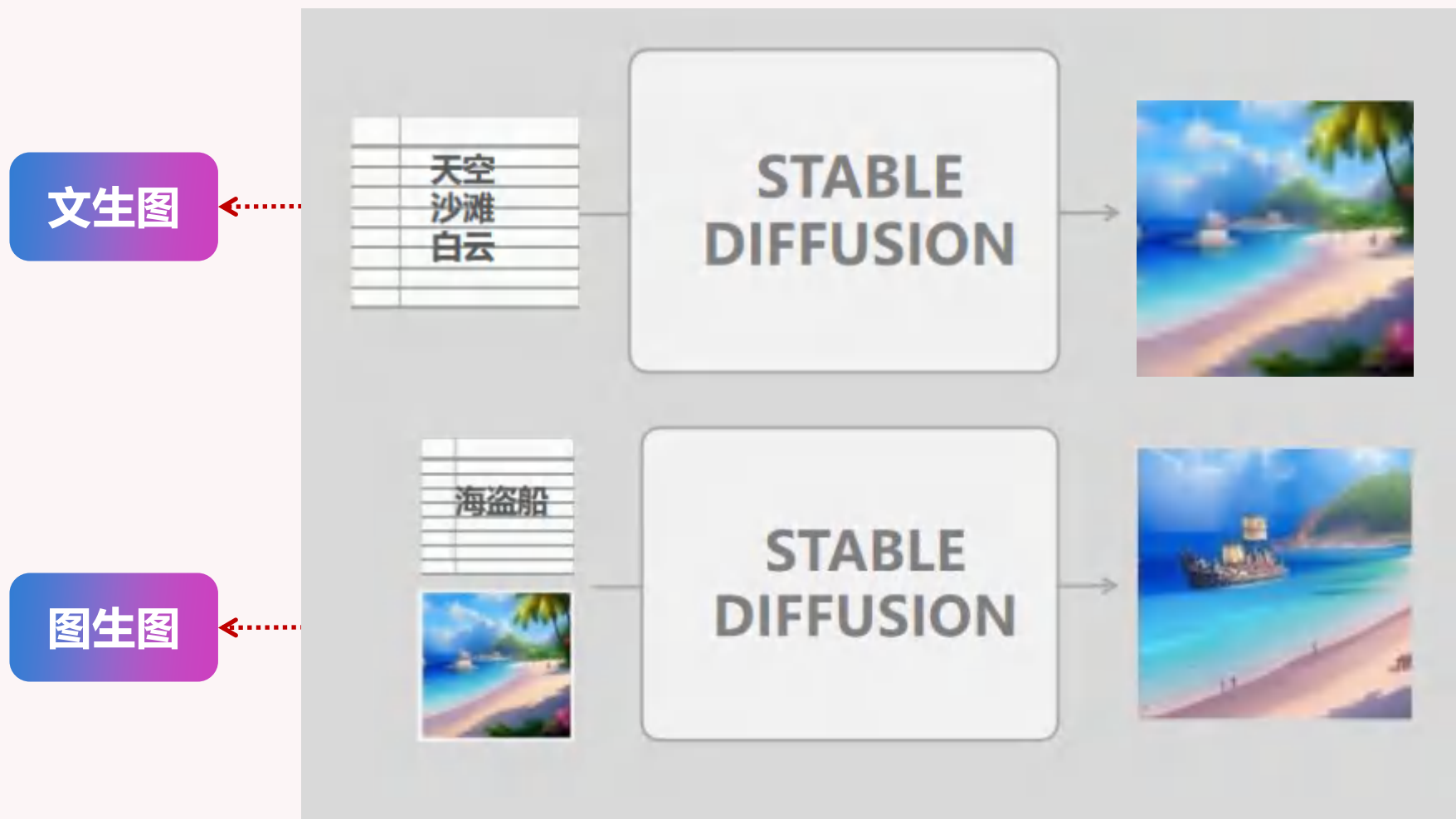
- 从回复中找到有价值的部分
- 判断结果质量：初步识别可靠信息
- 评估信息准确性：与其他来源进行对比
- 处理错误信息：指出错误，重新提问

多轮对话

- 追问
- 澄清
- 引导

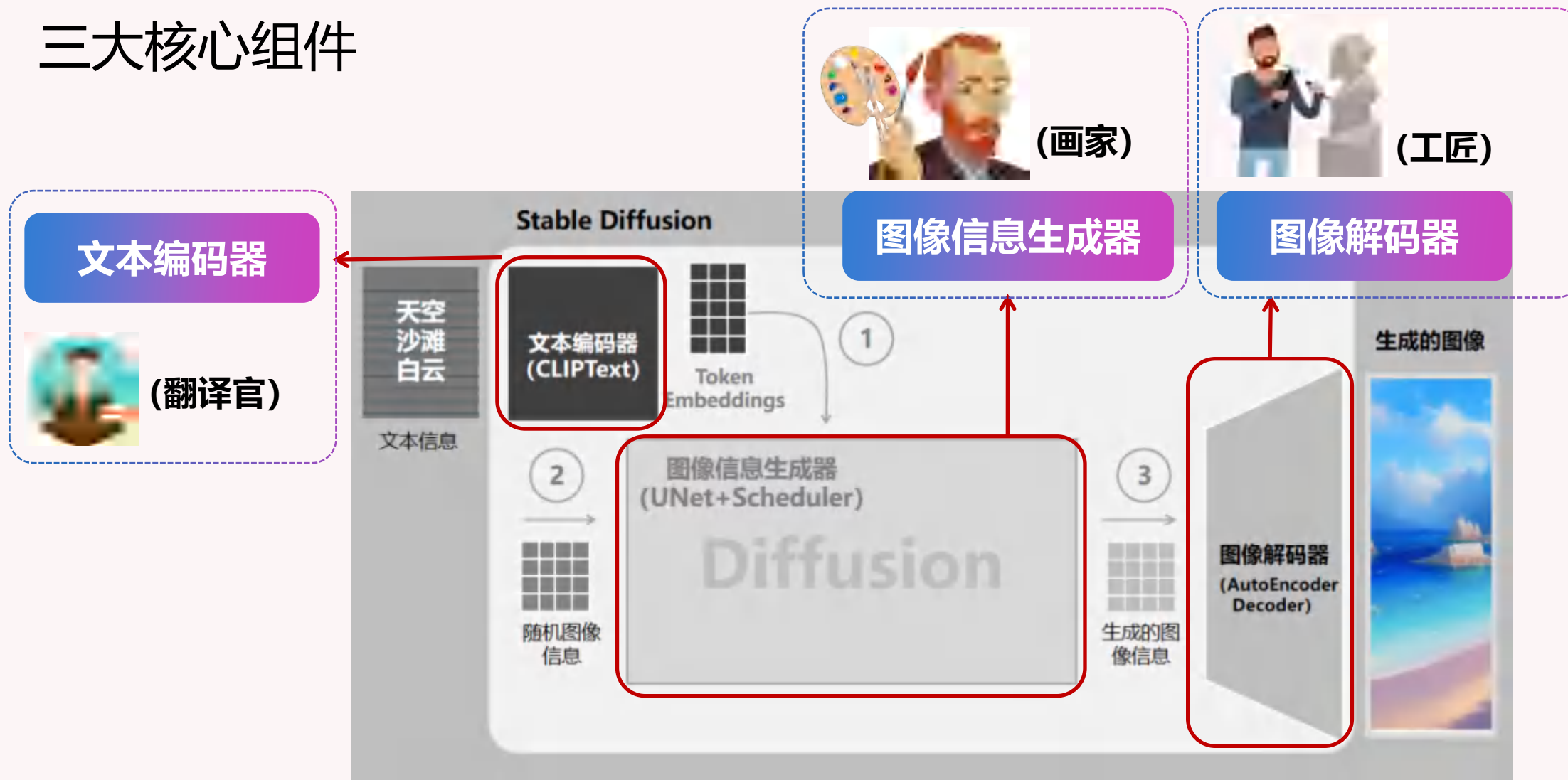
图像生成的奥秘

Stable Diffusion 模型如何实现图像生成?



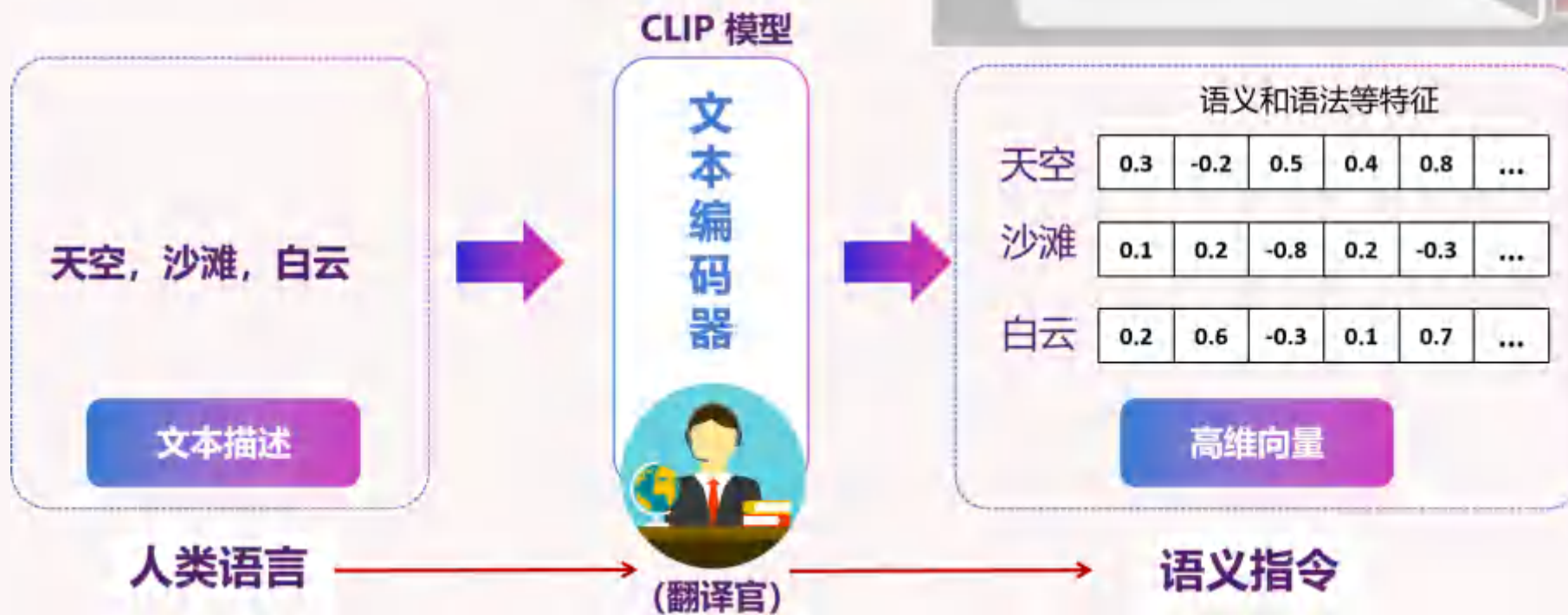
Stable Diffusion 模型的构成

三大核心组件



文本编码器

- 解析用户输入的提示词
- 将提示词转化为高维向量表示



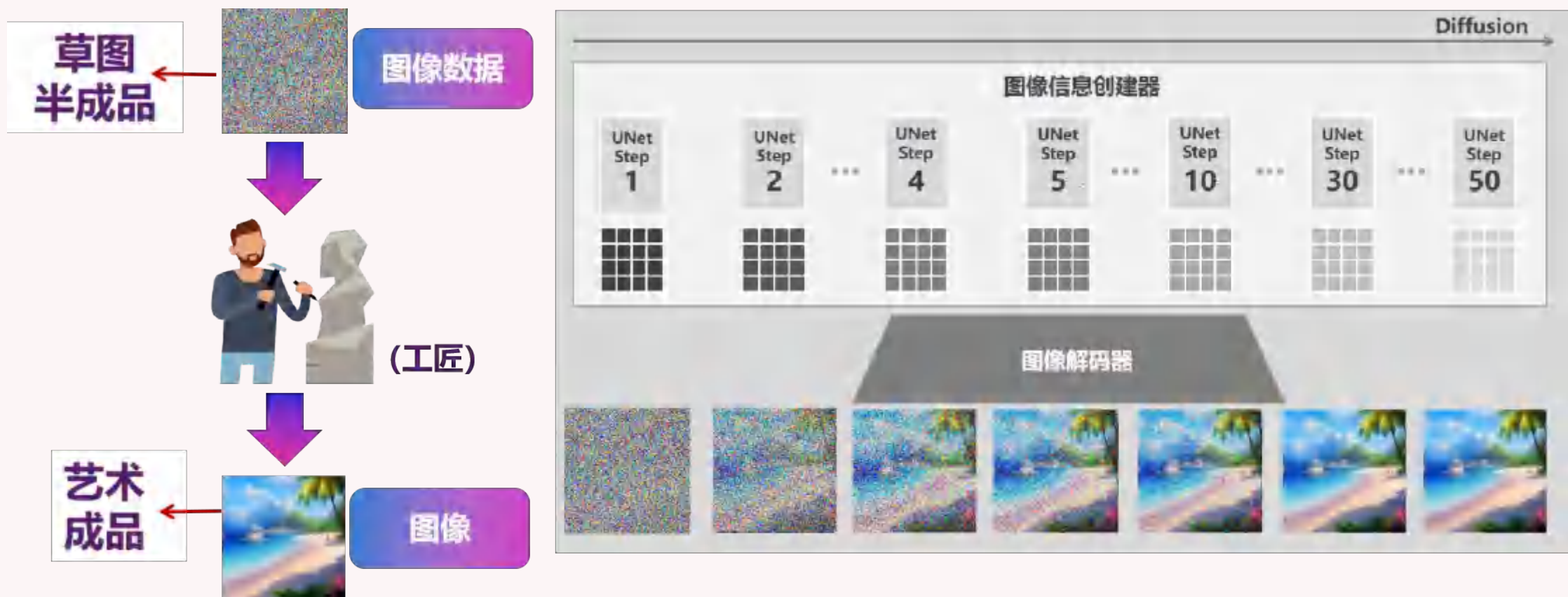
图像信息生成器

- 将输入信息逐步处理和转换为图像数据



图像解码器

- 接收来自图像信息生成器处理后的信息矩阵
- 将信息矩阵转换为可视化的图像



文生图的优势与局限

优势

- 降低门槛
- 提高效率
- 艺术风格多样化

局限

- 精确控制困难
- 随机性太强
- 复杂场景理解



PART 04 ▶

选择AIGC工具

本部分聚焦于如何科学选择和高效应用AIGC工具，以满足不同场景下的多样化需求。

首先介绍当前AIGC领域的工具类型，包括聊天对话机器人、图像生成工具、音频与视频生成工具以及搜索工具等，并列举了国内外代表性工具。

随后，以DeepSeek-R1、Kimi、豆包、腾讯元宝等具体工具为例，详细分析了它们的特点、优势及适用场景，展示了不同工具在推理能力、多模态支持、长文本处理和搜索效率等方面的差异。

最后，还提出了选择AIGC工具的依据，包括明确需求、评估工具性能和考虑使用成本等，帮助读者根据自身需求做出合理选择。通过深入分析和对比，本部分旨在为读者提供一份实用的AIGC工具选择指南，助力其在AIGC时代更好地应用人工智能技术，提升工作效率和创新能力。

Ai产品榜 · 国内增速榜

全球排名	Ai产品榜	产品名	网站(web)分类	12月上榜网站	12月上榜网站
			aicpb.com	Web访问量	变化
1		腾讯元宝	LLM	867.14K	351.12%
2		DeepSeek	AI ChatBots	11.01M	160.63%
3		AI 搜 百度	AI ChatBots	11.86M	93.04%
4		文心智能体平台 百度 Agents Platform	Agents Platform	617.86K	82.62%
5		豆包 抖音	AI ChatBots	31.77M	48.28%
6		即梦 AI 剪映	AI Image Generator	2.19M	45.50%
7		火山方舟	Model Training & Dep	3.49M	32.84%
8		Marscode	AI Code assistant	1.11M	29.34%
9		纳米AI搜索 纳米AI	AI Search Engine	359.43M	27.11%
10		知子	AI ChatBots	2.10M	23.79%
11		腾讯元宝	AI ChatBots	1.54M	20.68%
12		Kimi 月之暗面	AI ChatBots	38.37M	16.93%
13		墨刀AI	AI Design Tool	2.27M	13.95%
14		亿图脑图	AI Mind Map Generator	589.6K	11.83%
15		liblibai	Model Training & Dep	2.4M	10.71%

Ai产品榜 · 搜索引擎 全球

全球排名	Ai产品榜	产品名	网站(web)分类	12月上榜网站	12月上榜网站
			aicpb.com	Web访问量	变化
1		New Bing	AI Search Engine	1.87B	2.31%
2		纳米AI搜索 纳米AI	AI Search Engine	359.39M	27.10%
3		Perplexity AI	AI Search Engine	97.54M	9.55%
4		纳米AI搜索	AI Search Engine	8.32M	3.94%
5		You	AI Search Engine	5.54M	-12.30%
6		CSDN	AI Search Engine	4.06M	-2.03%
7		Primeya	AI Search Engine	4.64M	2.73%
8		1Ask AI	AI Search Engine	3.22M	1.41%
9		Kagi	Kagi Search	3.44M	15.61%
10		Consensus	AI Search Engine	3.37M	-0.45%
11		知子搜索	AI Search Engine	3.38M	8.40%
12		GenSwark	AI Search Engine	2.76M	28.90%
13		知子搜索	AI Search Engine	2.25M	5.49%
14		知子搜索	AI Search Engine	1.48M	1.14%

公众号：AI产品榜

AIGC工具

AI工具导航平台: <https://www.aigc.cn>

AI工具集: <https://ai-bot.cn>



工具类型	国内	国外
聊天对话机器人	DeepSeek、腾讯元宝、豆包、Kimi、智谱清言...	ChatGPT, Claude, X.com (Grok)
图像生成工具	通义万相、即梦、豆包	ChatGPT, Midjourney, Stable Diffusion
音频工具	音疯、网易天音、海绵音乐	Suno
视频生成工具	SkyReels、Vidu、可灵、海螺AI、即梦	Sora, Runway, Pika
搜索工具	秘塔、纳米AI、天工	Perplexity

AIGC工具现场演示

腾讯元宝 <https://yuanbao.tencent.com/>



AIGC工具现场演示

豆包 <https://www.doubao.com>

你好，豆包

让我们一起度过美好的一天!

AI 搜索

搜索最新资讯，让信息检索更高效

图像生成

按照你的想象，生成各种风格的图像

帮我写作

提供写作灵感，驾驭各种体裁和风格

阅读总结

速览长篇文章，迅速了解摘要和要点

试试这些例子

🔍 京东近期为整治“大企业病”都做了什么?

👤 一位半人半机器人的女孩，她通过电缆与一台巨大的计算机相连

📄 关于“端午出游”的小红书文案

📄 发布版infoQ大模型测评报告2024.pdf

帮我写作

图像生成

AI 搜索

PDF 问答

翻译

网页摘要

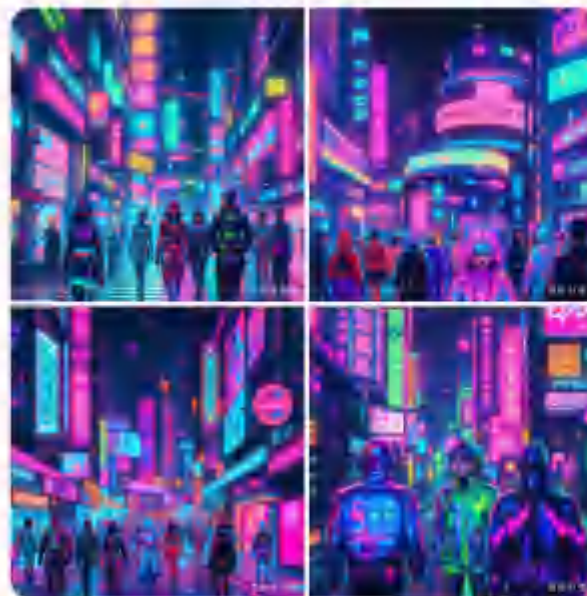
更多



请输入您的问题



赛博朋克风格的城市街头，霓虹灯闪烁，街上行人，探索奇物

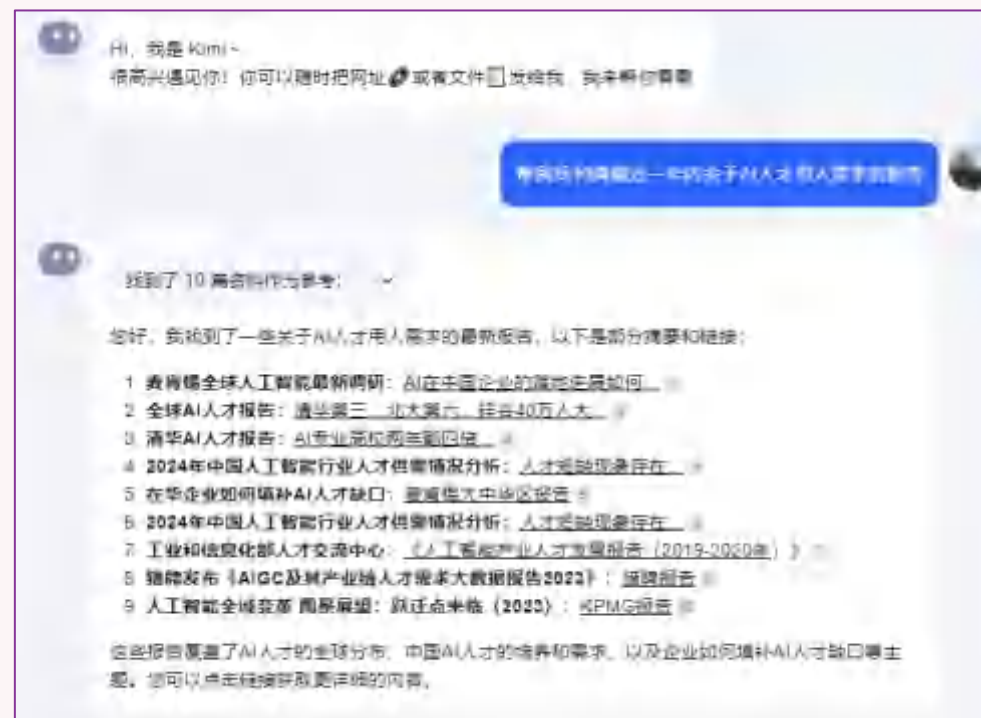


AIGC工具现场演示

Kimi <https://kimi.moonshot.cn/>



超长上下文
一次性阅读50份文档
支持20万字上下文输入



网络搜索
提高搜索效率
减少幻觉

AIGC工具现场演示



功能:

- 总结网页或PDF文档
- "去Kimi继续对话"



适用:

- 长视频内容 (B站/youtube) 的总结
- 新闻报道
- 行业报告



Kimi Copilot插件

AIGC工具现场演示

音乐工具：音疯 <https://www.yinfeng.cn/create>



AIGC工具现场演示

视频生成工具: Vidu <https://www.vidu.cn/>



一个3D形象的男孩, 穿着飞行夹克, 在公园滑滑板



AIGC工具现场演示

搜索工具：秘塔 <https://metaso.cn>



如何选择AIGC工具

1. 明确自身需求

2. 评测工具能力

3. 关注使用成本

AIGC工具	工具特点	大模型
DeepSeek	文本模态、推理能力强	DeepSeek-R1
豆包	多模态、语音情感能力强	火山
Kimi	多模态、超长文本、搜索和推理能力强	Kimi-1.5
智谱清言	多模态	GLM-Zero-Preview
通义千问	效率工具、代码能力强	Qwen2.5-Max
腾讯元宝	可以使用微信生态	接入：DeepSeek-R1

如何提升AIGC使用能力：持续更新自己的AI工具列表



1. 明确使用AI工具的需求和目标

2. 建立使用清单

- 文本类、绘图类、视频类、语音类、搜索类、设计类、专用工具类（办公类、笔记类、内容转化类...）、智能体类
- 以1~2个为主，其它为辅

3. 真实使用中测试和筛选

4. 保持2-3周的更新频次

AIGC技术正在重塑各个行业的未来，从内容创作到科学研究，从教育到商业应用，潜力无限。然而，技术的快速发展也带来了新的挑战，面对这些挑战，我们需要保持开放的心态，积极学习AIGC的基础知识，关注其在各行业的应用案例，跟踪最新发展趋势。

与AI共舞，实现AI与人类的完美融合。

让我们以DeepSeek-R1为起点，持续探索AIGC的无限可能。

在AI时代的技术浪潮中，我们既是见证者，更是参与者。



谢谢