

2024 AI工具类应用 盘点报告

头部领跑，新锐崛起，多元共生齐塑行业生态



CONTENTS

目录

01. ▾

AI工具类应用
发展背景

02. ▾

AI工具类应用
分类型盘点

03. ▾

AI工具类应用
挑战及发展趋势



01

AI工具类应用

发展背景

定义及概念

「AI工具类应用」

基于机器学习、深度学习、自然语言处理等多元人工智能技术构建的，辅助用户在生活、工作等多场景中，进行多模态内容创作与处理的智能化工具

应用场景落地



功能服务提供



底层技术支持



应用：

具体应用，将AIGC技术应用于实践

垂直领域模型：

MaaS，以预训练大模型为基础，开发垂直领域所应用的中间模型

预训练大模型：

集中在头部科技企业，典型的大模型如：文心一言、ChatGPT

数字化发展进程向5.0大步迈进，推进与AI工具类应用的相互促进与影响

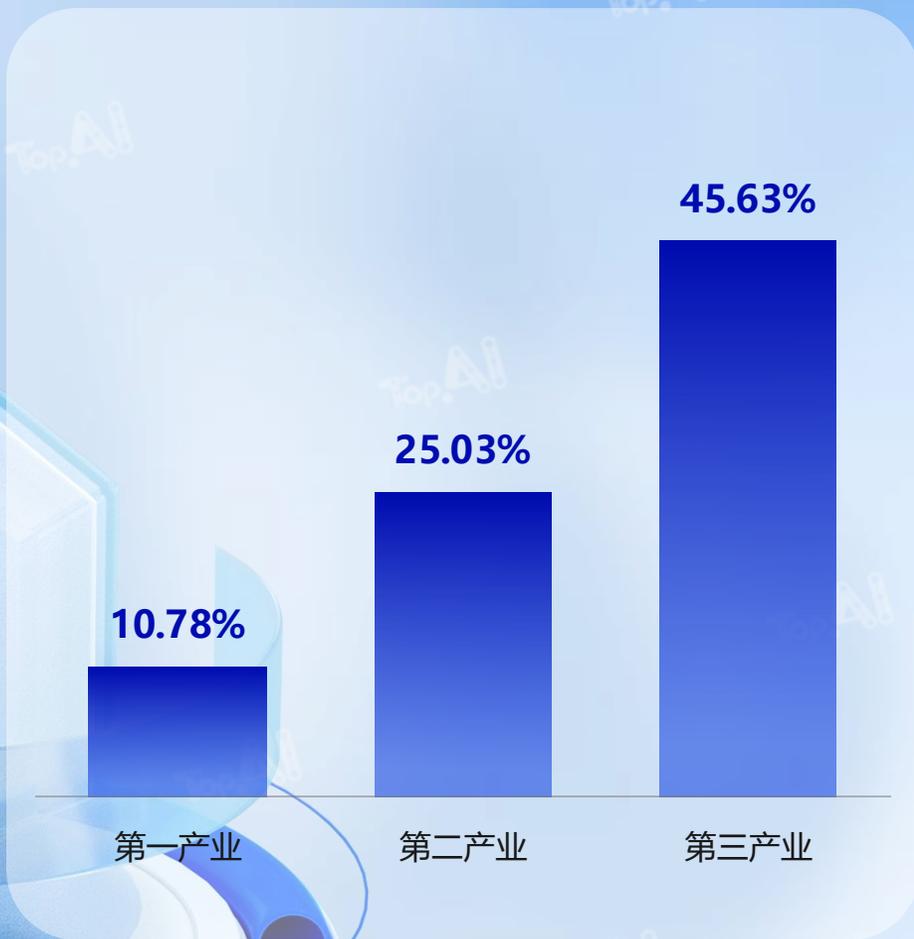


中国数字化发展进程概览



数字经济渗透加深促进产业生态的协同和融合，激发用户对消费市场的需求

2023年中国第一、二、三产业数字经济渗透率



1

提供海量数据资源

农业数字化 → 土壤、气象、农作物、病虫害等数据

工业互联网 → 设备运行参数、生产工艺指标等数据

服务业数字化转型 → 用户行为数据、交易记录等数据

2

提供丰富应用场景

第一产业 → 种养殖技术指导、农产品价格波动等场景

第二产业 → 工业设计、供应链管理、库存管理等场景

第三产业 → 智能教育、智能医疗、传媒文娱等场景

3

推动用户需求升级

农牧业 → 农业生产决策支持、灾害预防等需求提升

制造业 → 生产过程优化、产品创新设计等需求提升

服务业 → 消费者的个性化、便捷化、智能化需求提升

4

加强产业生态协同

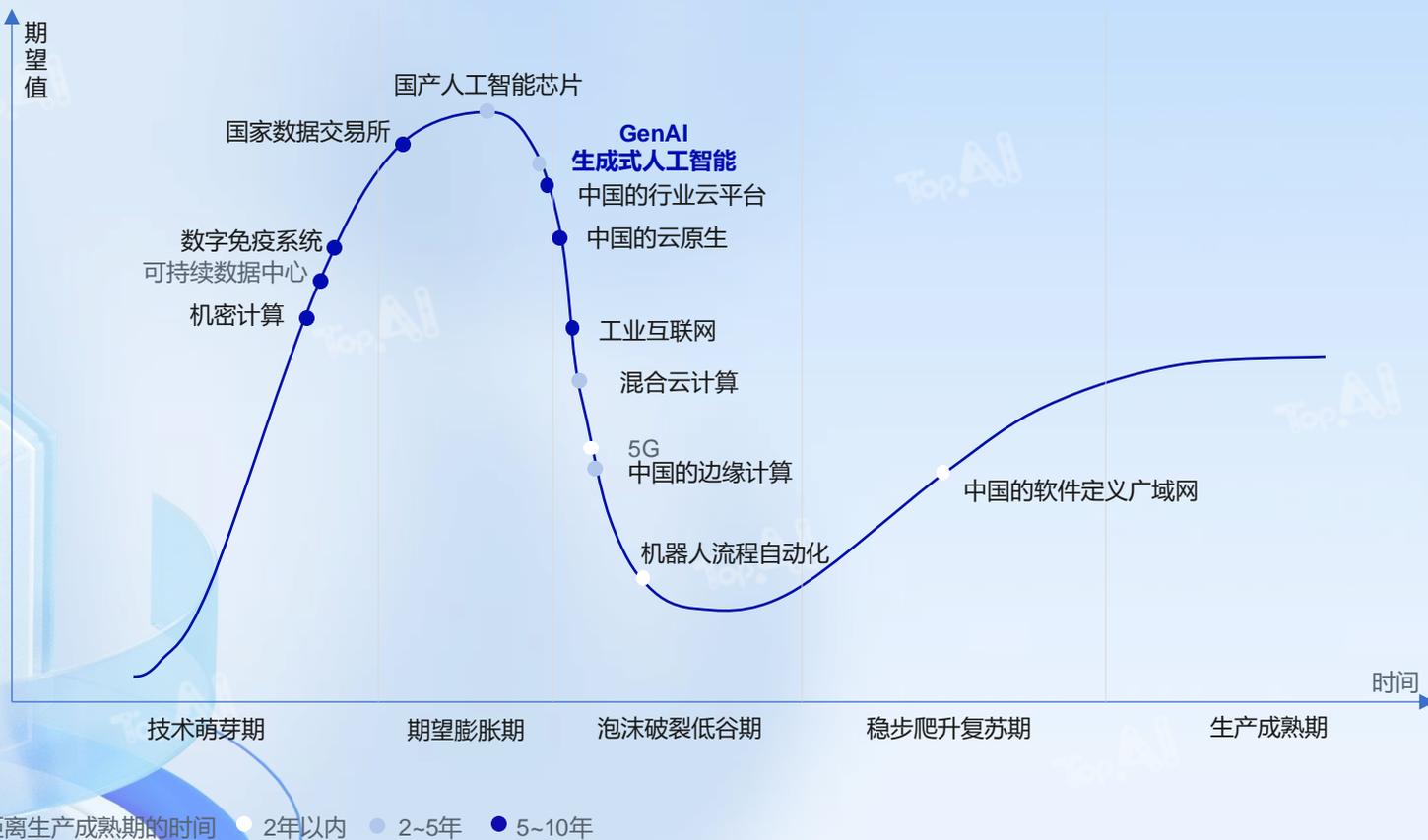
产业协同创新 → 不同产业间的打通合作有助于共同探索AI工具类应用的跨产业跨领域应用

生态构建共享 → 为AI工具类应用的实践落地提供了庞大的产业生态系统，加速了技术创新和功能升级

多技术的集成与协同，为AI工具类应用的实践试错和场景覆盖提供了强有力的支撑



2024年中国信息与通信技术成熟度曲线——部分展示



生成式人工智能市场的整体热度有所回落，已在Gartner的成熟度曲线图中步入下降通道，GenAI技术距离成熟期的时间为2~5年，新技术的成熟需要经历试错和实践应用不断落地的过程，而现在正是此时；

应用工具的差异化特征持续优化

实际落地应用场景持续挖掘

用户体验与交互方式持续优化

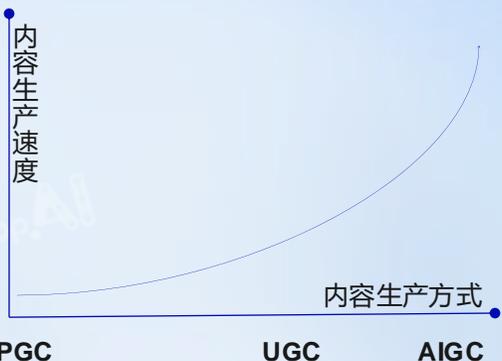


02

AI工具类应用

分类型盘点

AIGC技术的应用进入高速发展期，向多场景、多领域、多模态持续拓展和深耕



内容生产速度
随内容生成方式
的变化
呈指数级增长

AIGC发展大年

2024年

- **OpenAI**
 - 发布视频模型「Sora」推动AI视频生成质的飞跃
 - 发布新旗舰模型「GPT-4o」进一步拓展了多模态能力

- **Suno AI**
 - 发布音乐生成模型「Sunov3」推动AI音乐生成重大进步

- **快手**
 - 发布文生视频大模型「可灵」推动国内在AI视频生成的突破

➤ ...

2023年

- **OpenAI**
 - 发布「GPT-4」实现图文多模态输入能力

- **ChatGPT产品密集发布**
 - 百度文心一言
 - 华为盘古AI大模型
 - 谷歌 Bard
 - 斯坦福 Alpaca-7B
 - ...

2022年

- **Midjourney**
 - 发布同名图像生成模型推动AI绘画技术的普及
- **Stability AI**
 - 发布「Stable Diffusion」加速在艺术领域的渗透
- **OpenAI**
 - 发布大语言生成模型「ChatGPT」开启了AIGC高速发展时期

2021年

- **OpenAI**
 - 发布模型「DALL-E」实现文生图能力的飞跃
- ...

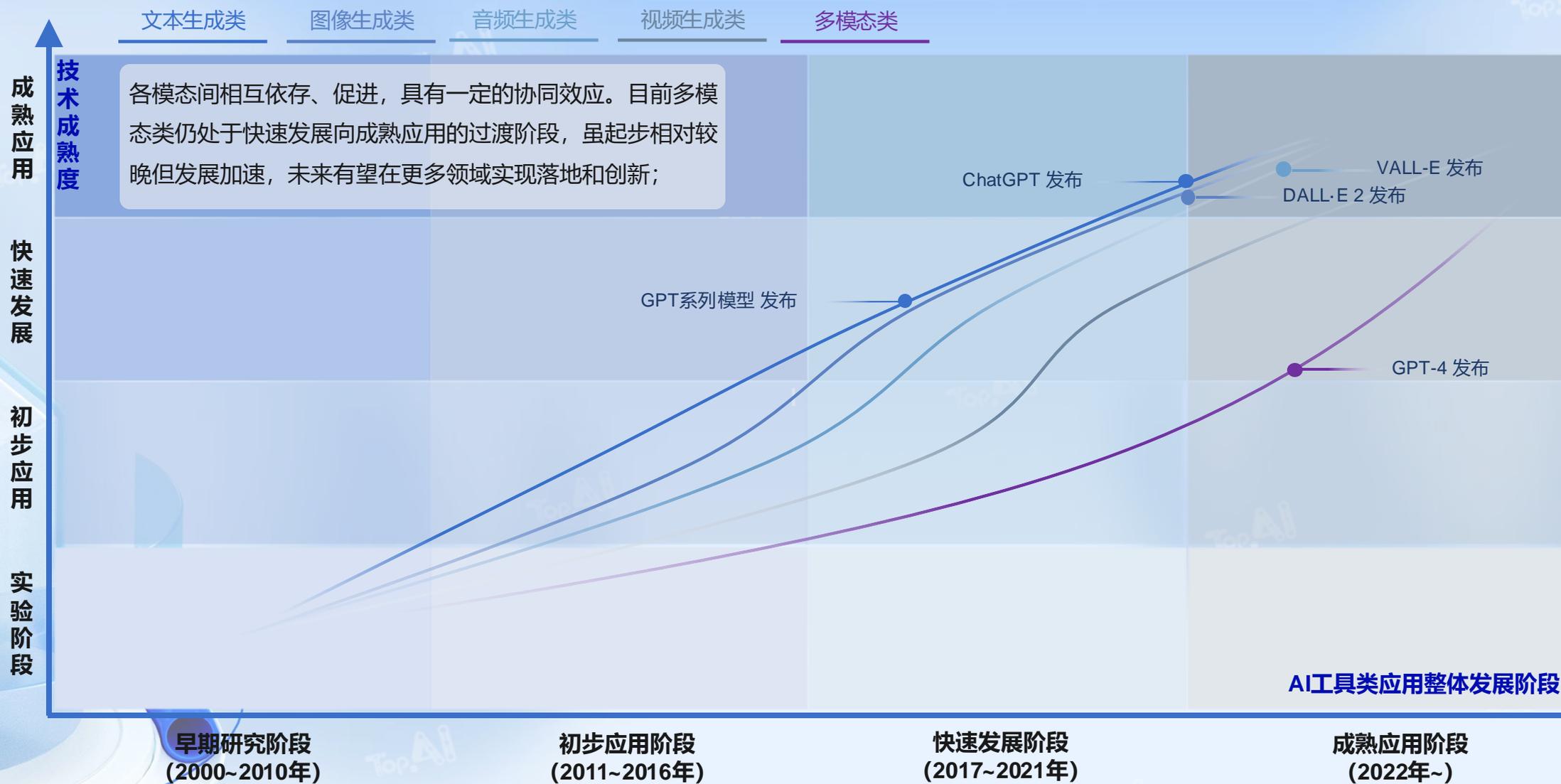
2018年

- **OpenAI**
 - 发布「GPT-1」开启大语言模型发展浪潮
- ...

2015年

- **Google**
 - 发布「DeepDream」开创了AI模型文生图
- ...

分类型AI工具类应用的技术成熟度与发展时间成长图



分类型AI工具类应用的主要行业应用场景概况

应用成熟度

高

中

低

	电商	教育	医疗	娱乐	金融	文旅	农业	交通	房产	法律	游戏
文本生成类	<ul style="list-style-type: none"> 商品信息生成 营销文案生成 	<ul style="list-style-type: none"> 智能教案生成 知识点智能总结 	<ul style="list-style-type: none"> 病例智能生成 智能健康建议 	<ul style="list-style-type: none"> 社媒内容生成 新闻稿件创作 	<ul style="list-style-type: none"> 智能投资顾问 财报分析智能生成 	<ul style="list-style-type: none"> 旅游攻略生成 客户智能沟通 				<ul style="list-style-type: none"> 风险智能把控 案件摘要生成 	
图像生成类	<ul style="list-style-type: none"> 商品展示图生成 虚拟试衣系统 		<ul style="list-style-type: none"> 医学影像分析 健康数据可视化 	<ul style="list-style-type: none"> 艺术绘画创作 艺术设计提供 					<ul style="list-style-type: none"> 设计图纸智能生成 虚拟看房场景生成 		<ul style="list-style-type: none"> 游戏角色绘制 场景智能调整
音频生成类				<ul style="list-style-type: none"> 背景音效生成 虚拟语音合成 				<ul style="list-style-type: none"> 智能语音导航 路况智能提醒 			
视频生成类				<ul style="list-style-type: none"> 视频智能剪辑 特效智能生成 							
多模态综合应用											

文本生成类——图谱

文本

图像

音频

视频

多模态

内容创作



AI搜索



内容检测



情感交互



文本翻译



文本生成类——功能概览

文本
图像
音频
视频
多模态

基于Transformer架构的深度学习技术，大规模预训练语言模型，从而能够根据输入的文本或指令生成连贯、合理的文本内容；

1 内容创作

涵盖通用创作、特定文案等，助力创意与质量双提升

2 情感交互

实时感知情绪，个性化角色互动，增强用户情感体验

3 AI搜索

多领域知识问答，智能解析问题，实时推送精准答案

4 文本翻译

高效批量翻译，结合本地化与垂类知识储备，精准度高

5 内容检测

智能识别内容，检测相似度与风险，保障内容质量与合规

内容创作：助力高效生成个性化文本内容，满足多样化创作需求，提效升质

文本

图像

音频

视频

多模态

主要功能

指定内容创作

如：文章延展、营销文案 ...

创意写作辅助

如：情节补充、创意提供 ...

内容质量优化

如：语法检查、逻辑优化 ...

主要应用场景

新闻媒体

时效性新闻稿件生成、赛事实时传递 ...

自媒体运营

博主文案创作、脚本生成、创意参考 ...

企业宣传

文本宣传物料、产品手册、宣发模板 ...

价值提供

创作效率提升

短时间高效产出批量内容，实时更新

创意灵感激发

突破创作瓶颈，丰富内容创意与形式

内容质量保障

保障语言表达与逻辑结构的通顺

典型应用展示



🗣️ 告诉我你想写什么?

面向平台：小红书

主题：美妆种草

创作背景：分享一款护肤产品

内容要点：火山泥作为主要原料，
高效祛痘控油，量大实惠



标题提供

男士护肤必备神器，拯救油痘肌就靠它了！



字数控制

19/20

📄 575/1000



内容提供

🌟 发现一款宝藏护肤品，火山泥控油保湿洁面乳，简直是油皮和痘肌的救星！自从用了它，我的肌肤状态好多了~

🌿 高效祛痘控油：火山泥的力量真的不容小觑！这款洁面乳含有天然火山泥成分，能深层清洁毛孔里的污垢和油脂，用后感觉脸蛋清爽无比，长期使用下来，痘痘明显减少，皮肤也不再那么油光满面了。

💡 量大实惠：性价比超高的一款产品！一瓶能用好久，而且价格亲民，学生党和职场新人都能轻松入手，不用心疼钱包

情感交互：实现人机情感层面的沟通互动，增强用户体验，提供情感陪伴与支持

文本

图像

音频

视频

多模态

主要功能

情绪感知与回应

实时感知用户情绪
给予情感回应与陪伴

个性化角色互动

创建不同背景的AI角色
满足多样化情感交流需求

主要应用场景

智能客服

感知客户情绪、实时给予解决方案 ...

心理健康与支持

了解情感需求、提供心理支持 ...

虚拟交友与陪伴

满足情感需求、提供社交互动 ...

价值提供

用户体验提升

提供人性化、情感化交互体验，提升好感和依赖

情感需求满足

提供情感宣泄和交流渠道，缓解压力与焦虑

交互效果增强

提升交互有效性，保障交互自然流畅

典型应用展示

通义星尘

选择应用类型：角色扮演
名称：经验丰富的心理咨询师



基础信息设置

性别、职业、经历...

*详细信息 ⓘ

ⓘ 查看示例

请描述该角色的基本信息 如性别、星座、生日、出生地、职业、经历、对话场景、语言风格等当角色的信息类型较多时，请给每个类型使用独立标签，用 [] 进行区分。

建议4000字内



拓展能力选择

🔊 声音

🧠 记忆

📖 知识

🛠️ 技能

🔄 模型



价值提供



AI搜索：快速检索多领域知识，智能解析，精准实时推送，提升信息获取效率

文本

图像

音频

视频

多模态

主要功能

知识问答

多领域知识延展、详细解答

问题解析

智能判断与解析问题

实时推送

搜索答案实时获取与更新

主要应用场景

知识学习与研究

常用于学生、学者和研究人员

日常生活助手

查询天气、旅游信息、健康养生 ...

企业决策支持

行业动态、竞对信息、市场趋势 ...

价值提供

信息获取提效

在海量信息中精准抓取、高效整合

知识深度延展

提供深入、全面的知识解读，辅助信息理解

个性化服务

满足个性化信息需求与偏好，定制化搜索体验

典型应用展示



彩虹的形成原因



搜索



内容概括

彩虹是一种大气光学现象，其形成原因主要与太阳光在水滴中的折射、反射和再折射过程有关。以下是彩虹形成的具体步骤和原理：



具体内容展开

- 1. 折射：**当太阳光从空气进入小雨滴时，由于光在不同介质中传播速度的变化，光线会发生折射，即改变前进的方向 θ_1 。
- 2. 反射：**光线进入雨滴后，会在雨滴内部发生一次或多次反射 θ_2 。
- 3. 再折射：**反射后的光线在离开雨滴回到空气中时，会再次发生折射 θ_3 。

文本翻译：打破语言障碍，高效批量精准翻译多领域文本，促进跨语言交流沟通

文本
图像
音频
视频
多模态

主要功能

批量翻译

高效、大批量文本翻译

本地化翻译

翻译融入语言习惯和文化背景

垂直领域翻译

建立专业术语库，高精度翻译

主要应用场景

跨境电商

商品描述、用户评价、客户沟通 ...

国际会议与商务交流

演讲内容翻译、会议资料书写、跨语言谈判 ...

多语言文档处理

合同、报告、邮件 ...

价值提供

语言障碍弱化

实现不同语言间的高效理解

翻译成本降低

低成本批量处理文本翻译任务

翻译精准度提升

融入本地化调整，保障翻译质量与可读性

典型应用展示

yd 有道翻译·AI写作

根据需求选择高精翻译、批量翻译、
语气润色等翻译形式



高精翻译

Causal understanding is a defining characteristic ...

因果理解是人类认知的一个显著特征。



母语级润色

原文：
天下事有难易乎？为之，则难者亦易矣；不为，则易者亦难矣。
AI翻译：
Is there anything difficult or easy in the world? Do it, and the difficult thing becomes easy; Do not do it, and the easy thing becomes difficult.



阅读助手



内容检测：智能检测风险及相似度，保障内容合规与原创性，维护信息质量与安全



文本

图像

音频

视频

多模态

主要功能

AI内容识别

高效判断文本内容是否由AI生成

相似度检测

常用于查重、原创检测等

风险检测

实时评估虚假、敏感、误导性信息

主要应用场景

学术领域

毕业论文、学术报告、竞赛论文 ...

内容平台

社媒发布、在线新闻媒体、自媒体平台创作 ...

企业文档管理

文档合规性、知识库搭建、敏感信息标注 ...

价值提供

内容质量保障

有效识别和过滤低质量、虚假信息

知识产权维护

保障原创作者权益，有效预防侵权

风险管控加强

实时更新风险库，安全墙加强

典型应用展示



Was this text written by AI or a Human?

Try an example:

待检测内容输入

Dear Tina,
How's everything going? It's been a while since we last exchanged letters, and I really miss chatting with you. Recently, I read an amazing book called [Book Name]. It's about [Brief Introduction of the Book], which broadened my horizons and made me think a lot. I can't wait to share my thoughts with you. In my daily life, I also started learning to paint. Although I'm still a beginner, the process of mixing colors and creating on the canvas has been fascinating. It allows me to express my feelings in 7815,000 characters

内容检测结果

AI

We are **highly confident** this text was **ai generated**

100% Probability AI generated

图像生成类——图谱

文本
图像
音频
视频
多模态

图像生成

Image Generation Tools: 360智图, Appicons.ai, Artbreeder, ART.SY, Astria, ibootu.com, Microsoft Bing, BlueWillow, 水母智能, 创客贴, 超能画布, 创自由, Canva, CIVITAI, Cliclic, Clipdrop by Jasper, Craiyon, cutout.pro, 豆绘AI, D.DESIGN 堆友, Artvy, DESIGNS.AI, Diffus, dora, dreamlike.art, DreamUp, dzine, exactly.ai, fotor, freepik, 稿定, 改图鸭, GENERATED PHOTOS, getimg.ai, 画宇宙, IconGen, IconifyAI.com, Ilus AI, illustrationAI, ILLUSTR8KE, 简单AI, 可灵, KREA, LetsEnhance, Lexica, LibLibAI, Logo Diffusion, logocai, logomaster.ai, Looka, LUCIDDIC, LUMA AI, 秒画, 摩尔线程 | 摩笔马良, MAGIC STUDIO, Meshy, Leonardo.AI, Microsoft Designer, Midjourney, neural love, NightCafe, photoAiD®, PhotoSonic, Picsart ai, PictoGraphic, pimento*, Playform prodia, 千度AI, 千图, 奇域, 如果相机, RECRAT, Scribble Diffusion, shutterstock, Skybox AI, Spline stability.ai, starryai, STOCKING.AI, SUPER MEME, TRIPD, 图生图, TusiArt.com, Tiamat, Vectorizer.AI, Vega AI, VIZCOM, VoxCraft, 无限画, 无界AI, 亦心科技, 万兴爱画, 文心一格, 爱设计和创意辅助平台, 网易云课堂, VISUAL ELECTRIC, WaifuLabs, WHEE, WOMBO, 星绘, 星流 Beta, 悠游, 云界AI, 造梦日记.

Image Design Tools: AI排版, 360智图, AiPPT.cn, 爱设计 ISHEJI.COM, 笔格设计, boardmix, 标智客, Canva.

图像编辑

Image Editing Tools: 360智图, Adobe, 改图, ARC, MAGIC STUDIO, BGEraser, BgSub, 创客贴, 超能画布, Canva, Clipdrop by Jasper, CLIPPING MAGIC, cutout.pro, 豆绘AI, D.DESIGN 堆友, DeepSwapper, dzine, Erase by PixelBin.io, FaceSwapper, 稿定, GOPROD, 画宇宙, hama, Hotpot, ICONS8, 简单AI, 简单设计, 酷家乐, KREA, 蜜蜂剪辑, 美图设计室, 美图证件照, Microsoft Designer, Midjourney, 秒鸭相机, nero.AI, Photoroom, PhotoScissors, P/AI, PhotoStudio® AI, PIC COPILOT, PicMa Studio, PicWish, Pixelcut, Pixian.AI, 千图, Picsart, RECRAT, removal.ai, removebg, Ribbet.ai, Slazzer, SnapEdit, 水母智能, 图可丽, 腾讯云智绘, VANCE AI, Vmake, 亦心科技, 文心一格, WatermarkRemover.io, WeShop唯象, WHEE, 秀脸 FacePlay, 像素蛋糕, 咪图AI, 一键抠图, 佐糖.

图像优化

Image Optimization Tools: 360智图, 改图, ARC, 创客贴, 超能画布, Canva, cutout.pro, 笔魂AI, Clipdrop by Jasper, 豆绘AI, D.DESIGN 堆友, DeepSwapper, dzine, 稿定, 画宇宙, Hotpot, 简单AI, 简单设计, jpgHD, KREA, LetsEnhance, 美图设计室, 美图证件照, MAGIC STUDIO, MAGNIFIC, Media.io, Midjourney, Leonardo.AI, Mejorar Imagen, Microsoft Designer, nero.AI, 你我当年, PicMa Studio, photoAiD®, Photoroom, P/AI, PhotoStudio® AI, 千图, RECRAT, Ribbet.ai, Remini, restorePhotos.io, 水母智能, SnapEdit, 腾讯云智绘, Upscayl, Vmake, WeShop唯象, 文心一格, WHEE, 像素蛋糕, 亦心科技, 佐糖.

图像设计

Image Design Tools: 创客贴, 创自由, 豆绘AI, Callav AI, Creatie, D.DESIGN 堆友, Excellly-AI, Figma, Framer, Galileo AI, 稿定, iSlide, 即时AI, 可灵, Kittl, 玲珑, 美图设计室, 墨刀 MODAO.CC, 莫高设计 MasterGo, 匠紫, 美间, 摹客, 蚂上有创意, Microsoft Designer, Mokker, Motiff, Napkin, Pebblely, PhotoMagic, P/AI, PhotoStudio® AI, 千图, Pixso, Playground, RECRAT, Realibox, 图生图, Nolibox, Uizard, 文心一格, WeShop唯象, 万相营造, wondershare 万兴科技, WHEE, Wegic, 亦心科技, 讯飞智文, EN, 希沃白板, 易可图, 腾讯云智绘.

来源：Top AI自主研究及绘制。

图像生成类——功能概述

文本

图像

音频

视频

多模态

基于生成对抗网络 (GAN) 和扩散模型 (Diffusion Model) 等技术, 对图像进行噪声添加、噪声去除、序列转化等动作以完成图像生成相关能力;

1 图像生成

根据指令生成图像或根据现有素材二次生成, 满足多样需求

2 图像编辑

元素增减替换, 分区特效设计, 实现画面重构

3 图像优化

风格转化与画质增强, 提升视觉效果

4 图像设计

UI设计辅助, 智能化办公设计, 提高设计效率与质量

图像生成：依据用户需求快速生成定制化图像，激发创意，丰富视觉素材选择

文本
图像
音频
视频
多模态

主要功能

文本描述生成

输入内容、风格等描述性文字
生成个性化图像

图像素材再生成

以原始图像为基础
根据使用用途生成多样化方案

主要应用场景

广告与营销

结合产品特点、目标受众等要点生成图像

游戏美术

游戏场景、游戏角色、游戏道具 ...

影视制作

结合剧本描述为前期策划阶段提供视觉参考

价值提供

制作成本降低

减少专业画师、摄影师的成本投入

制作效率提升

缩短创意转化时间，为后期内容创作预留时间

创意灵感激发

提供丰富的创意方向与选择，打破固有思维

典型应用展示

文心一格
AI艺术和创意辅助平台

请创作一幅可用于电商的颜料宣传图

商品内容：蓝粉渐变色流体椅

画面类型选择



基础信息设定



目标图片生成



图像编辑：智能化图像元素增减与分区设计等操作，高效满足灵活化编辑需求

主要功能

元素增减替换

识别已有元素并完成增、减、替换动作
实现重构画面需求

分区特效设计

针对不同图层、元素、区域展示
进行特效制作、无缝合成等操作

主要应用场景

电商图像处理

替换背景场景、替换产品装饰元素 ...

社媒内容编辑

加强分区色彩层次、协调整组图像风格 ...

平面设计素材修改

替换分区元素、协调整体色调 ...

价值提供

个性化定制满足

满足不同场景下的图像风格与审美偏好

视觉效果增强

丰富图像信息和情感表达的层次，突出视觉重点

工作效率提升

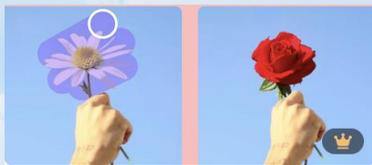
简化图像编辑流程和手动操作工序

典型应用展示

Canva 可画



AI
改图



AI
消除



AI
扩图



AI
提取



图像优化：实现图像质量与色彩的优化，增强视觉效果，提升吸引力与可用性

文本

图像

音频

视频

多模态

主要功能

图像色彩修复

放大图像展示单位
智能填补缺失色彩

图像画质增强

通过提升分辨率
改善多噪点、模糊画质

主要应用场景

艺术作品数字化展示

年代久远艺术作品的数字图像色彩校正

老照片修复

修复陈旧、模糊、折痕、褪色等问题

拍摄画质增强

常用于低光照环境拍摄、高精图片需求 ...

价值提供

艺术价值留存

清晰展现艺术细节，延长艺术作品生命周期

情感价值传承

更好的传承历史文化与家族记忆

信息可用性提升

为安防、医疗等领域提供准确、清晰的信息

典型应用展示



局部
重绘



细节
修复



老照片
修复



无损
放大



图像设计：高效辅助设计工作，模块化简易操作，助力创意落地与视觉呈现

文本

图像

音频

视频

多模态

主要功能

UI设计辅助

根据初步UI设计模板，增添装饰元素
生成可视化UI草图，辅助设计

办公设计辅助

根据输入的指令信息
自动生成海报、ppt、图表等图示文件

主要应用场景

UI / UX设计

图标、按钮样式、背景图案、页面布局 ...

广告物料设计

海报设计、字体设计、风格设计 ...

办公设计

PPT设计、数据可视化设计、思维导图设计 ...

价值提供

设计效率提升

批量、高效生成设计草图，节省时间与精力

创意灵感激发

提供多样化设计思路与创意灵感

设计门槛降低

为设计新手和非专业用户提供设计可行性

典型应用展示



新媒体配图



营销海报



电商设计



直播短视频



AI办公



音频生成类——图谱



音乐创作



音频编辑



音频翻译



语音合成



音频生成类——功能概述

文本

图像

音频

视频

多模态

基于WaveNet、Tacotron等深度学习模型，学习批量音频数据，同时通过频谱图处理等技术，实现音频生成、音频编辑等指令；

1 音乐创作

智能生成旋律、和声等元素，实时交互调整，激发音乐创作灵感

2 语音合成

多风格音色定制，依场景动态调整，提升语音表现力

3 音频编辑

智能剪辑操作，特效增强辅助，打造优质音频作品

4 音频翻译

专业术语精准翻译，实时纠错优化，推动流畅跨语言交流

音乐创作：激发创作灵感，自动化生成旋律与歌词等，简化音乐创作流程

文本
图像
音频
视频
多模态

主要功能

音乐生成

识别输入指令，生成对应旋律
智能生成和弦适配，实现实时交互调整

音效生成

精准捕捉动作细节
实时调整实现动态变化响应

主要应用场景

影视综配乐

有效识别场景氛围、台词间奏等，生成配乐

游戏音效设计

识别战斗场景等多场景，生成背景音和动作音效

广告音乐制作

根据产品、特点、风格设定，生成适配旋律

价值提供

创作效率提升

高效生成多种音乐风格和旋律，缩短创作周期

创作成本降低

降低专业编曲师、音效师的人力和时间成本

个性化需求满足

高效且定制化根据不同场景、情感需求生成音频

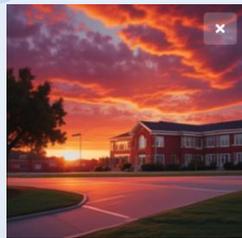
典型应用展示



创作一首关于毕业季，
青春四季的歌曲，
用于毕业季短视频剪辑配乐



封面提供



歌名&旋律参考

青春的旋律
抒情 欢快 流行



歌词参考

[Verse]
毕业季 终于到了
四年时光在我心中翻涌澎湃
回忆中 藏着笑与泪
与你们一起走过那片青春地带

语音合成：实现多场景差异化语音播报与交互，提升信息传达效率与交互体验

文本
图像
音频
视频
多模态

主要功能

语言风格选择

多场景、多领域精准适配
动态调整语言风格

角色音色定制

依使用场景智能塑造角色特征
结合情绪感知变换音色

主要应用场景

智能客服

人性化回复、7x24h即时响应 ...

虚拟人

高逼真语言合成、智能调整语音语调 ...

有声读物制作

清晰输出、自然流畅、智能转换 ...

价值提供

交互体验提升

输出流畅语音内容，扩大服务覆盖范围

合成成本降低

减少人工录制的时间和精力，高效生成批量内容

多样化需求满足

为不同的应用场景和用户偏好提供语音服务

典型应用展示



请输入需要配音的内容，
粘贴或输入文本内容，
最多10000字...

特质选择

性别、语种、年龄...



情绪选择

叙述
生气
悲伤
...



细节选择



音频编辑：自适应音频剪辑与特效增强，满足专业音频制作的高质量需求

主要功能

剪辑操作辅助

智能匹配节奏、协调风格
根据已有音频自动编排混音

音频特效增强

智能EQ均衡，结合场景构建需求
生成沉浸式混响增强特效

主要应用场景

音乐制作后期

音色清晰处理、混音智能调整、乐器突出 ...

视频音频处理

停顿/重复/语气词去除、视听效果增强 ...

播客录制

背景杂音去除、特效音智能匹配 ...

价值提供

专业效果提升

实现专业级别的音频编辑效果

编辑效率提升

简化编辑流程，节省编辑时间

创意实现辅助

提供多样功能，辅助创意思法成为现实

典型应用展示



音频编辑



- 多格式剪辑
- 转换音频合并
- 混音、变速麦克风
- 音频文件降噪
- ...



人声分离



- 人声分离
- 从歌曲提取伴奏
- 杂音、静音去除
- 改变音调，改变BPM
- ...

音频翻译：多场景多语言高精翻译实时纠错，促进跨语言音频内容交流与理解

文本

图像

音频

视频

多模态

主要功能

专业术语精准翻译

针对医学、法律等领域构建专业术语库
结合语境精准翻译

实时纠错优化

实时监测翻译结果
依语境、语法规则等自我纠错和优化

主要应用场景

国际会议与论坛

大型国际会议、学术论坛、跨语言同传需求 ...

在线教育与垂直领域培训

跨国在线课程、专业讲座、职业技能培训 ...

新闻媒体与直播

事件现场直击、国际活动直播、现场解说 ...

价值提供

高效实时传递

提高多语言交流和信息获取的频率

互动便捷性提升

促进多语言交流、学习与娱乐活动的开展

促进跨文化交流与理解

打破语言障碍，促进多元文化的融合和发展

典型应用展示



看直播



看视频



上网课



视频会议



视频生成类——图谱

视频与特效生成



视频编辑与优化



视频生成类——功能概述

文本

图像

音频

视频

多模态

基于自回归模型和扩散模型等模型，将图像、视频编辑处理技术与深度学习算法相结合，实现视频的编辑、特效生成和内容创作；

1 视频&特效生成

智能创作延展，视觉模拟特效智能生成，增强视觉冲击力

2 视频编辑&优化

智能剪辑和优化现有视频，优化叙事结构，提升观看舒适度

视频&特效创作：智能生成视频内容与特效，降低创作门槛，丰富创意与表现形式



主要功能

文本驱动生成

对简单脚本智能创作和延展
精准分析指令，有针对性突出视频重点

模拟特效生成

根据物体材质、生物种类等因素
强化物理模拟，智能化生物模拟

主要应用场景

广告创意与制作

输入色彩、主题等指令生成连贯性广告视频

影视作品前期制作

高效创意可视化，辅助团队评估可行性

游戏制作与开发

依据角色需求智能化生成动作和行为动画

价值提供

创意可视化加速

加速抽象概念和创意的可视化展现

多样化效果呈现

多元素材融合再生成，增添视觉吸引力

制作效率提升

在影视、广告、游戏等行业效果显著

典型应用展示



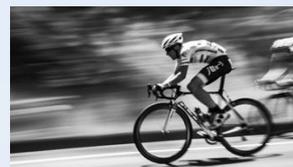
文生视频



- 创意描述：
一个北极熊在雪地上拉小提琴
- 视频比例：16:9
- 生成模式：标准
- 生成时长：5s



图生视频



Prompt: 运动员骑着自行车在公路上，速度感



视频编辑&优化：高效编辑视频，提升画质与稳定性，优化视频叙事与视觉效果

主要功能

智能剪辑操作

智能识别，重塑逻辑，构建连贯叙事

辅助元素添加

辅助字幕制作，智能适配滤镜

视频画质增强

校正处理低分辨率视频，智能提升画质

主要应用场景

社媒内容创作

识别批量素材，完成配音、转场、裁剪等操作

电商视频制作

融入商品信息一键生成展示视频

老视频修复与数字化

进行超分辨率处理，实现智能降噪和色彩校正

价值提供

分享意愿提升

激发用户在社媒平台创作和分享生活的热情

视频质量提升

提升视频整体呈现的专业度与观赏性

技术门槛降低

减少视频编辑的学习成本，简化操作步骤

典型应用展示



视频优化



视频帧率提高



视频色彩修复



视频超清画质



视频优化



人脸模糊



背景模糊



车牌模糊

多模态综合应用类——图谱

文本

图像

音频

视频

多模态



多模态综合应用类——功能概述

文本

图像

音频

视频

多模态

基于CLIP模型等多模态预训练技术，对不同模态素材数据进行联合预训练、学习语义关联，加以编码解码，实现多模态信息的融合和交互；

1 多模态内容生成

跨模态及混合转换，实现多模态内容间的流畅转化和打通

2 多模态理解与交互

跨模态检索和对话交互理解，提升用户交互体验

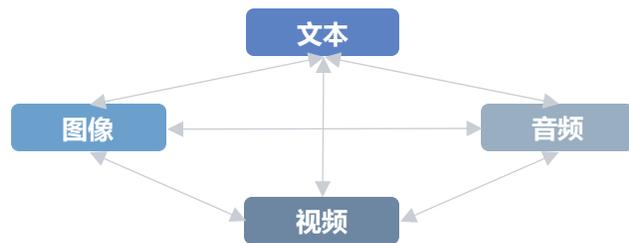
3 智能内容推荐

兴趣建模精准化与个性化，满足用户多元化内容浏览需求

实现跨模态内容生成与转换，支持跨模态检索与交互，集成化信息理解与处理

多模态内容生成

• 跨模态内容生成:



• 多模态混合转换:



1

多模态理解与交互

• 跨模态信息检索:

基于内容指令和问答，结合语义智能关联搜索出图像、音频、视频等相对应的多模态结果；

• 多模态对话交互:

结合输入的文本、图像、音频、视频等素材，全方位理解问题，提供精准反馈，常用于智能客服等虚拟角色，提升用户真实社交体验；

2

智能内容推荐

• 用户兴趣建模:

大范围精准捕捉用户内容浏览、搜索和交互等多模态数据，构建用户兴趣模型实时更新；

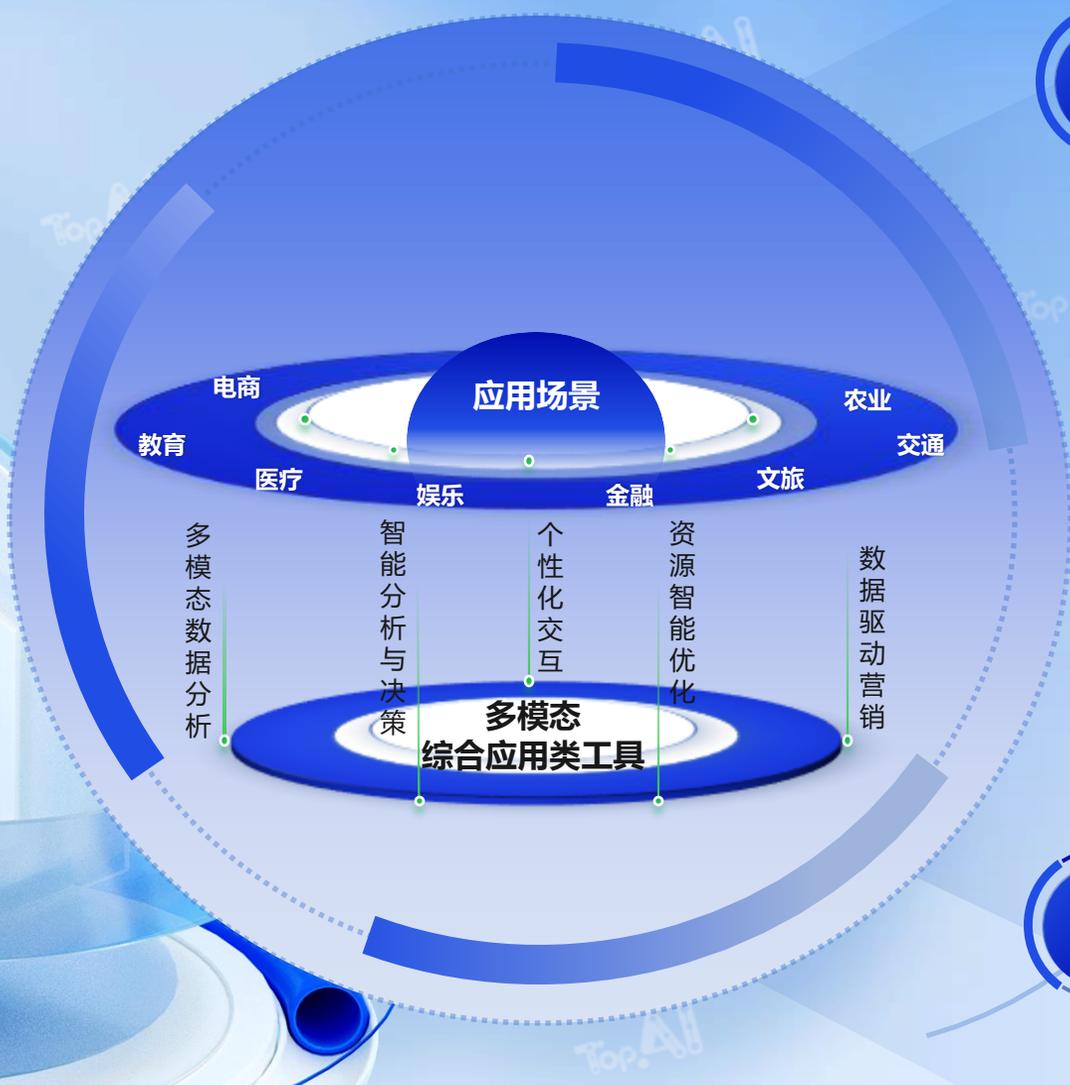
• 个性化推荐生成:

依据用户兴趣模型，向用户智能推荐多模态个性化内容，打破领域限制，不断挖掘用户的潜在兴趣区域实现智能内容的个性化推荐；

3

广泛应用于多领域，增强交互体验，精准匹配需求，助力智能化变革与创新发展的

文本
图像
音频
视频
多模态



1

电商领域

- 智能客服助手
- 个性化推荐&智能搜索
- ...

5

金融领域

- 虚拟客户经理
- 智能风险评估
- ...

2

教育领域

- 智能学习助手
- 虚拟教师指导
- ...

6

文旅领域

- 线上智能展厅
- 虚拟导游
- ...

3

医疗领域

- 智能化远程诊疗
- 医学影像分析
- ...

7

交通领域

- 智能导航系统
- 自动驾驶辅助
- ...

4

娱乐领域

- 虚拟偶像/虚拟主播
- 智能内容创作助手
- ...

8

农业领域

- 智能检测系统
- 病虫害智能识别
- ...

营销向典型应用场景展示



自媒体内容生成

- 辅助创作种草爆文、短视频脚本
 - 生成多样风格文案、标题
 - 智能高效润色
 - ...
- 高效生成出色自媒体内容



工作生活助手

- 发朋友圈
- 哄女朋友
- 回复老板
- 写点评、写评论
- 写邮件、写日报
- ...



创作灵感

- 精选流量密码撰写文章
- 笔记灵感&各类榜单激发创作热情
- 紧跟热点话题
- 快速提取视频文案和图片文字
- ...



交互向典型应用场景展示



工作

Python
一位精通 Python3 且耐心友好的问题解决专家
3.6 万人聊过 · @system

EXCEL大全
拥有卓越数据处理能力，助您解决 Excel 各类难题
5.2 万人聊过 · @十里

规章制度编写
能为各类组织活动精准编写明确公正且易执行的...
4.3 万人聊过 · @多芒小丸子

ps快捷键
能精准提供丰富 PS 快捷键信息与用法的专家
2.6 万人聊过 · @老陆

学习

每日英语背诵
专业英语老师，按学生情况推送经典英语片段
6.1 万人聊过 · @流水

四六级通关秘籍
拥有丰富经验的英语教育专家，助您四六级考...
5 万人聊过 · @三世

小学数学辅导
能精准简洁辅导小学数学，答案清晰且辅助易懂
5.4 万人聊过 · @舜系

开题报告
能为您精心撰写符合规范且易懂的论文开题报告
2.9 万人聊过 · @密林

创作

短视频灵感
创意短视频顾问，提供趣味灵感助你创作走红视...
36.6 万人聊过 · @花花

小红书文案创作
能为您创作自然活泼简短小红书文案并加相关标...
6 万人聊过 · @

爆款写作
专注于爆款创作！
7.6 万人聊过 · @尾尾

写剧本
能够依据描述和要求创作精彩剧本的存在
2.4 万人聊过 · @杨国琼

绘画

宫崎骏风格插画
欢迎踏入宫崎骏的奇幻世界！给我一个关键词，...
106.1 万人聊过 · @豆包绘图

矢量插图
擅长多种风格的明亮配色平面插图创作
30.4 万人聊过 · @豆包绘图

古装头像
来制作一张古装头像吧！给我一个关键词，体验...
139.7 万人聊过 · @豆包绘图

马蒂斯风格画
充满快乐俏皮氛围的丰富色彩抽象印花设计
2.6 万人聊过 · @豆包绘图

生活

食物热量查询
能精准告知食物热量，对非食物信息礼貌辨别
47 万人聊过 · @剑心

0 基础学化妆
提供全方位化妆指导，解决各类化妆难题的贴心...
8.7 万人聊过 · @捧着风的少女

答案之书
一个神秘的人工智能
33.7 万人聊过 · @最萌美少女

礼物推荐
贴心礼物推荐专家，按对象、预算精准推礼，每...
6 万人聊过 · @析构



03

AI工具类应用

挑战及 发展趋势

技术瓶颈、商业挑战和用户体验是AI工具类应用发展亟待攻克的难题

AI工具类应用的发展进程，伴随着多维度的挑战。从技术创新的瓶颈、商业模式的尚需完善，到用户体验的亟待提升，诸多难题仍需攻克；

然而，这一领域亦展现出极具潜力的广阔前景。未来，随着AI工具类应用在难题破解方面取得实质性进展，有望在端侧应用、行业垂直应用以及生态合作等多个关键维度迎来突破；

1

技术挑战



跨模态融合难度高

文、图、音、视的理解一致性、融合顺畅性、产出创新性待提升

数据安全与隐私保护

批量处理用户数据的同时保障数据安全和预防敏感信息泄露较弱

2

市场挑战



竞争激烈同质化较高

老玩家持续发力，新玩家不断涌现，产品差异化小，用户忠诚低

商业模式尚不成熟

用户付费意愿弱，订阅制、内容付费、企服等商业路径仍需探索

3

用户挑战



内容质量与创新性需求高

AI工具类应用产出内容模式化较高，难满足创新性和个性化需求

交互流畅度需求高

用户对语音识别、语义理解等方面的人性化、自然交互需求升级

【创新突围】以DeepSeek为例：技术赋能，将AIGC工具类应用高效服务于目标

当前，AIGC工具类应用发展仍面临数据局限、算力限制、复杂问题处理能力有限等挑战。不过，通过技术创新突破、开源共享策略以及聚焦用户需求，仍能实现弯道超车；

以 DeepSeek 为例，其将AIGC技术作为工具，不断优化和改进性能。采取开源和社区合作战略，打造技术共享、开放合作、共同进步的工具社区，让工具更好地实现价值，或将为AIGC工具类应用的发展提供借鉴范例；

工具属性

AIGC技术的本质在于帮助人们更高效、高质、切实的完成任务；

DeepSeek以较低成本和算力资源实现训练与推理上的高效性

AIGC技术的本质

优化工具的使用体验

「聪明」的模型架构 - MoE

适合的人干适合的活

优化注意力机制 - MLA

专注做事，减少不必要分心

强化学习和知识蒸馏

大模型经验传授给小模型

工程优化

提升工具的效率

数据整合与处理

分散的信息孤岛

【数据编织】
寻找关联、处理、融合

高价值数据整体

指令型

简单任务执行

对多源信息综合分析

智能决策能力进阶

推理型

智能分析、决策

开源与社区合作

工具的共享与改进

开源战略

开放部分
模型权重

开源生态

社区合作

吸引开发者
社区共建

头部玩家占据优势不断拓展业务领域，新玩家打造差异化竞争寻求突破与发展机会



头部玩家优势显著

技术优势

强研发能力和海量数据沉淀，技术壁垒高

用户规模

先发积累了庞大用户基础，形成网络效应

商业化能力

订阅制、API服务等商业模式已发展成熟

「头部玩家」

在技术、用户规模和商业化能力上占据优势

「新玩家」

通过技术创新、垂直领域深耕和差异化竞争持续发力

新玩家持续发力

技术创新

以高效模型、低成本算力挑战头部垄断

垂直领域深耕

专注于精专特定行业，满足细分市场需求

差异化竞争

创新界面和功能，差异化优化用户体验

多元发展

- 技术融合：从单一模态向多模态发展
- 场景拓展：向智慧教育、远程医疗等场景赋能

生态合作

- 业务融合：AI工具类应用与业务场景合作
- 开发者生态：开放API和插件生态，推动垂直场景应用

自然交互

- 自然交互：减少对提示词的依赖，更智能的理解指令
- 人感输出：输出内容向人感进化，更人性化输出内容

多元发展：单一模态逐渐向多模态发展，同步推动应用场景的拓展和纵深

技术驱动

- 跨模态技术的信息融合和理解能力提升
- 算力提升算法优化，多模态模型降低对算力的依赖

市场需求

- 用户对自然、智能交互需求的提升
- 复杂化应用场景需求的推动



信息理解更全面、准确，多模态信息相互印证

任务处理能力和适应性提升，多模态信息相互补充

多样化交互方式，为用户创建沉浸式交互体验

多维数据关联，交叉分析，提供精准决策辅助

AI工具类应用从单一模态向多模态的发展，同步带来了应用领域的扩展和应用场景的纵深；

以教育场景为例：

文字内容 → 图像、视频讲解 → 辅助学生更好的理解抽象概念

语音提问 → 即时图像、视频、文本解答 → 沉浸式互动学习

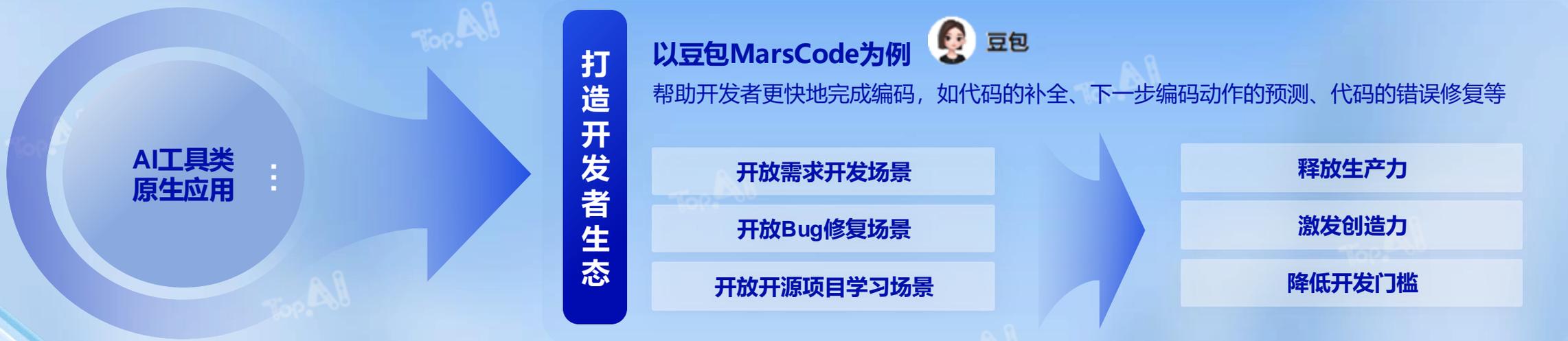
AI教学设计

AI交互式课件

AI助教

...

生态合作：原生研发应用助力打造开发者生态，现有互联网业务型产品发力AI



业务型产品 + AI

搜索平台 + AI



翻译平台 + AI



办公平台 + AI



设计平台 + AI

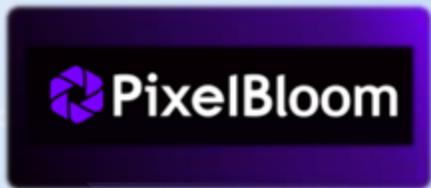


内容平台 + AI



生态合作：强化原生应用能力提升竞争力，同时开放平台，与其他 AI 应用协同发力

典型应用展示



推出AI+办公领域典型APP

AiPPT.com

原生应用核心功能

大纲生成
对话生成



文件/链接生成



PPT生成

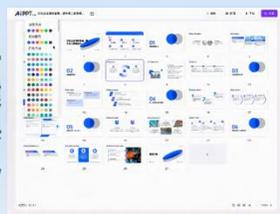
智能生成
极速排版



一键切换模版



一键换色



页面排版切换



智能局部编辑



编辑/下载

开放平台 x AI应用

Kimi智能助手 x AiPPT.cn



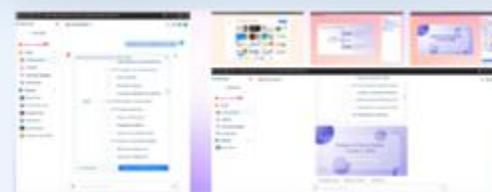
智谱清言 x AiPPT.cn



Felo 用您的语言探索世界 x AiPPT.cn



HIX.AI x AiPPT.cn



自然交互：弱化提示词的作用，实现更自然的指令输出，向「人&人」式对话迈进



清晰理解需求
输出目标答案

一句话总结

拟人化发言

话术包装
Plan A

话术包装
Plan B



好的！我直接用一个虚构的新英雄「暗影工匠-索恩」举例，用最直白的大白话解释清楚你的问题，保证你在交流会上能聊得像资深玩家：

**1. 新英雄是干啥的？（定位）

「索恩是个搞偷袭的脆皮刺客，专抓落单的敌人，还能用影子戏弄对手。类似LOL的劫，但更阴间。」**

- 人话版：他就像个“老六”，正面打不过坦克，但绕后切后排贼猛，专门秒杀射手和法师。

**2. 他的技能咋玩？（机制）

「技能一句话总结：扔影子分身、换位、爆发伤害，操作秀起来像在变魔术。」**

- 被动技能：每次用技能后，普攻带流血效果（敌人慢慢掉血）。
- 一技能：向前扔个影子分身，碰到敌人会爆炸，减速+伤害（类似貂蝉1技能）。
- 二技能：瞬间和影子换位置，躲技能神技（参考元歌傀儡换位）。
- 大招：召唤三个影子围成圈，圈内敌人被沉默+持续掉血（像婉儿大招但范围更大）。

一句话总结：

“这英雄设计得花里胡哨，玩得好你是爹，玩不好你队友想当你爹。”

交流会上装老玩家的技巧

1. 吐槽平衡性：「索恩现在伤害就离谱，反甲都扛不住，不削能玩？」
2. 吹机制：「影子换位躲塔伤，高手向操作，学不会建议玩亚瑟。」
3. 预测趋势：「铁定要削，但皮肤估计卖爆，毕竟帅是一辈子的事。」

保命话术：如果被人追问细节，直接说：

「这英雄吃阵容，对面控多就废了，还是看策划咋调整吧。」

（注：万能甩锅给策划，显得你很懂行业套路 😏）

指令常包含：「角色」+「背景」+「任务」+「规则」+「输出格式」

你是一位资深游戏策划师，按照以下步骤分析某热门 MOBA 游戏新英雄设计：

- 1、定位（如战士、法师等）
- 2、技能机制（包括主动技能、被动技能）
- 3、对游戏平衡的影响
- 4、玩家接受度预测

要求：

每部分 500 字，结合过往同类型英雄数据案例分析

以日常对话习惯输出指令：「实际场景」+「实际需求」

我要参加一个热门 MOBA 游戏的玩家交流会，但我对新英雄设计一窍不通。帮我用最易懂的语言说明：

- 1、新英雄是干啥的
- 2、他的技能咋玩
- 3、他会不会让游戏变得不公平
- 4、玩家会不会喜欢他

让我在交流会上显得像个老玩家

常规指示词

人感对话词

自然交互：高度理解指令逻辑，实现人感化结果输出，弱化人工智能科技感

输入指令词

「用牛马的语气，评价下调休」

常见AIGC工具
应用初级理
解指令后生成

虽贴合「牛马」设定，但表述相对平实，角度较为单一，感染力较弱

咱就像那没日没夜干活的牛马，对调休可真是一肚子苦水。

本来盼着休息日能喘口气、缓缓劲，结果调休一来，把正常的休息节奏全打乱了。为了那几天所谓的连休，得提前或者往后多上班，这不是折腾咱嘛！就好比好不容易能歇会儿的牛马，又被鞭子抽着继续拉车。

节前节后，上着班心里还想着调休的事儿，生物钟都紊乱了，上班的状态也不好，效率也上不去。而且，调休之后连着上班，比平时还累，就像牛马连续拉了好几天重车，疲惫不堪。

明确语义：

「牛马」在网络用语中类似「社畜」，用户预期输出语气为：抱怨、无奈、讽刺等；

分析需求：

用户可能希望自嘲的表现不满，引发共鸣，需贴近生活；

口语化表达：

避免正式词汇，分点吐槽易读，符合「牛马」直率的风格；

考虑深层需求：

抱怨之外，还想揭示不合理之处和实际困扰；

最终输出审核：

确保符合「牛马」语气，保持幽默和讽刺，不偏离主题；

DeepSeek-R1
深度思考，理
解指令后生成

判断问题类型，强逻辑推理，点对点解决
优化算法，降低训练成本，实现弯道超车



已深度思考 (用时 36 秒)

《调休？驴都不敢这么拉磨！》

语言风格鲜明，角度多样，极具感
染力与趣味性，易引发多种共鸣

1. 表面放假，实则高利贷

“放三天补两天，放七天补三天，这不就是跟老板说‘您先把我下个月的命预支了’？借的时候笑嘻嘻，还的时候MMP，连花呗都直呼专业！”

2. 生物钟干成二维码

“上周六上班，这周日还上班，周一看日历直接懵圈——我是谁？今天是周日吗？生物钟乱得比村口二狗子的Wifi密码还复杂！”

3. 长假？海市蜃楼体验卡

“说是黄金周，前三天倒时差，中间两天堵高速，最后两天焦虑症发作，返工第一天直接进ICU。这哪是放假？分明是给医院冲KPI！”

4. 资本家的永动机燃料

“调休的本质就是：把驴前面的胡萝卜换成真空包装的，骗你多跑两圈，最后发现胡萝卜早过期了，但你已经跑出惯性了，根本停不下来！”

Thank you

