



Research and
Development Center

百尺竿头须进步，十方世界是全身

AI 行业 2024 年度策略报告

冯翠婷 传媒互联网及海外首席分析师
执业编号：S1500522010001
邮 箱： fengcuiting@cindasc.com

刘旺 传媒互联网及海外团队研究助理
邮 箱： liuwang@cindasc.com

证券研究报告

行业研究

行业投资策略

AI 行业

投资评级 看好

冯翠婷 传媒互联网及海外首席分析师

执业编号：S1500522010001

邮箱：fengcuiting@cindasc.com

刘旺 传媒互联网及海外团队研究助理

邮箱：liuwang@cindasc.com

信达证券股份有限公司

CINDA SECURITIES CO., LTD

北京市西城区闹市口大街9号院1号楼

邮编：100031

AI 行业 2024 年度策略报告：百尺竿头须进步，十方世界是全身

2024 年 01 月 21 日

- **2022 年底，AIGC 热潮在全球范围内兴起，人工智能的主题行情开始出现。**2022 年 11 月 30 日，OpenAI 的全新对话式 AI 模型 ChatGPT 正式发布，人工智能板块开始掀起热潮。2023 年我们认为这是人工智能的元年，年初 Meta、百度等纷纷推出大模型，AI 大厂英伟达在芯片设计中加入 AI，标志着计算硬件的进步，欧洲的 AI 立法和中国的监管措施反映了全球对标准化 AI 治理的转变。这一年，AI 开发者们逐渐认识到，人工智能不仅仅是一个概念或远景，而是一场正在发生的、深刻影响各行各业生产力的历史性技术革命。
- **全球大型语言模型行业快速发展，创新不断，OpenAI 的 GPT 系列引领趋势。**中国科技巨头也积极参与，百度、腾讯、阿里云、昆仑万维、科大讯飞、商汤科技、360 等企业，大中小企业齐头并进推动行业进步。截至 2023 年 10 月底，国内已经发布 238 个大模型，通过国家网信办备案的大模型超过 20 个，大模型已进入规模落地应用阶段。据赛迪研究院数据，2023 年我国大语言模型市场规模将达到 132.3 亿元，增长率达 110%。尽管国内模型还在早期阶段，但市场潜力大。23 年 7 月网信办发布《生成式人工智能服务管理暂行办法》，也表明监管支持态度，展现出与全球同步的发展潜力。
- **AI 大模型高速迭代，多模态 AI 应用持续探索迭代。**相比于纯文本大语言模型，多模态模型也有自己的基座模型、预训练模型以及带有涌现能力的模型，如在基础模型 CLIP、DALLE 的基础上也已衍生出类似 Flamingo、PaLM-E 等视觉多模态语言模型，通用大语言模型在高速迭代中。从 AI 应用类型来看，目前大语言模型应用不只局限于 NLP 文本生成，多模态大模型已经可以处理来自不同模态（如视频、图像、语音、文本等）的多种信息，如自然语言处理、多模态推理、计算机视觉和音频处理等，实现模态之间的转换和统一。各行业也正期待着在多模态大模型方向上能有类似“ChatGPT 时刻”的出现。
- **投资建议：**受益于全球生成式 AI 持续发展，文生图/视频/音乐/代码等多方面应用的迭代更新有望长期给诸多行业带来降本增效。建议关注，AI 大模型端：腾讯、百度、阿里巴巴、昆仑万维、科大讯飞等；应用端：**1) AI+办公：**万兴科技、福昕软件、金山办公；**2) AI+游戏：**紫天科技（AI 游戏先行者，率先发力互动剧）、昆仑万维（新推出 AI Agent 开发平台，天工大模型六大矩阵看点丰富）、巨人网络（《原始征途》表现优异，休闲品类有望受益新渠道）、掌趣科技（与悠米、行者 AI、layabox 等合作布局 AI+游戏领域）、神州泰岳（游戏出海表现稳健，非游戏板块拥抱 AI）、盛天网络（游戏业务稳定基本盘，带带电竞及给麦

带来增量贡献)、恺英网络、三七互娱、吉比特、宝通科技、名臣健康、姚记科技、顺网科技、汤姆猫、星辉娱乐等；3) **AI+影视**：上海电影（集团支持+IP 优势）、华策影视（影视剧储备丰富+微短剧布局）、光线传媒（动画电影受益于文生视频进展）等；4) **AI+出版**：果麦文化、南方传媒、中南传媒、凤凰传媒、皖新传媒等；5) **AI+教育**：世纪天鸿等；6) **AI+营销**：三人行、芒果超媒、蓝色光标、分众传媒、易点天下、兆讯传媒、浙文互联、遥望科技等。7) **IP 资源**：奥飞娱乐、中文在线。8) **美股标的**：微软、谷歌、Adobe、Unity 等；9) **港股标的**：腾讯、美图公司、拼多多、阜博集团、柠萌影视、小米集团、小鹏汽车等。

➤ **风险因素**：生成式 AI 公司资本开支不及预期、AI 应用迭代不及预期

一、2023 人工智能领域的里程碑，全球 AI 技术的飞速进步与治理挑战并进.....	5
二、全球大模型风起云涌，中国 AI 行业正在崛起.....	7
三、生成式 AI 应用及展望	11
四、投资建议.....	19
五、风险因素	20
分析师声明	22

表目录

表 1: 2023 年 AI 大事记.....	5
表 2: 国外大模型简介	8
表 3: 国内大模型简介	9
表 4: 不同应用场景 AI 模型访问量排名.....	11

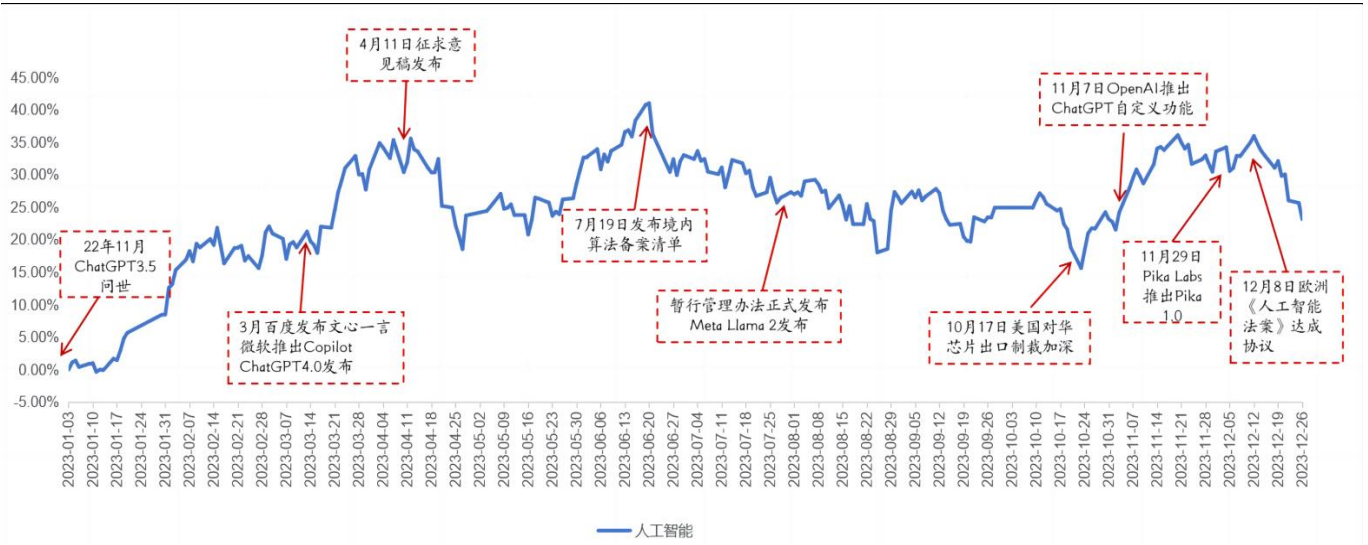
图目录

图 1: A 股人工智能板块 2023 年涨跌幅.....	5
图 2: 全球大模型发展历程	7
图 3: 国内外大模型发布情况.....	9
图 4: 文本和视觉多模态大模型的发展轨迹对比	15
图 5: DALL·E3 较 DALL·E2 在图片丰富度、细节方面的进步对比	16
图 6: Midjourney V5.2、V6 生成图像细节对比	16
图 7: Runway 视频剪辑工具功能	16
图 8: Runway Gen-2 模型	17
图 9: Pika 文生视频内容	17
图 10: 图/文生成 3D 人物模型过程示意	18
图 11: GPT-4V 对指向性提示以及逻辑思考能力	18
图 12: GPT-4V 对视频理解和时间序列的理解能力	19

一、2023 人工智能领域的里程碑，全球 AI 技术的飞速进步与治理挑战并进

2022 年底，AIGC 热潮在全球范围内兴起，人工智能的板块行情开始出现。2023 年可以说是人工智能板块行情的元年，也是 LLM（大语言模型）的元年。OpenAI 通过 GPT-3.5 系列大型语音模型微调而成的全新对话式 AI 模型 ChatGPT 正式发布，人工智能板块开始掀起热潮。Meta 推出了新型 AI 语言模型 LLaMA-13B，OpenAI 发布了 GPT-4 并不断降低成本，百度推出了文心一言；应用方面微软推出了基于 GPT-4 的 Microsoft 365 Copilot，很大程度上改变了办公应用的生产力，算力方面，大厂英伟达的 GPU 不断迭代，标志着计算硬件的进步，GPU 的供不应求导致英伟达业绩的持续超预期。监管方面，各地政府也是迅速调整，意大利对 ChatGPT 的临时禁令及其后的重新启用凸显了对 AI 治理的迫切需求，欧洲的 AI 立法和中国的监管措施反映了全球对标准化 AI 治理的转变。这一年，AI 开发者们逐渐认识到，人工智能不仅仅是一个概念或远景，而是一场正在发生的、深刻影响各行各业生产力的历史性技术革命。

图 1：A 股人工智能板块 2023 年涨跌幅



资料来源：同花顺，信达证券研发中心

表 1：2023 年 AI 大事记

日期	事件
2023/1	在 2022 年 12 月 ChatGPT 发布五天突破百万用户之后，2023 年 1 月 ChatGPT 月活用户达到了 1 亿
2023/2/25	Facebook 母公司 Meta 宣布推出新的、有人工智能（AI）支持的大型语言模型（LLM）LLaMA-13B，LLaMA 模型集合的参数从 70 亿到 650 亿个不等，较小的 AI 模型可能帮助催生能在个人电脑（PC）或智能手机等设备上本地运行、类似 ChatGPT 风格的语音助手。
2023/3/15	人工智能初创公司 OpenAI 正式公布最新一代人工智能语言模型 GPT-4，GPT-4 的诞生是 OpenAI 在放大深度学习方面的努力的最新里程碑，GPT-4 是一个多模态大模型（接受图像和文本输入，生成文本），可以更准确地解决难题，具有更广泛的常识和解决问题的能力，更具创造性和协作性。
2023/3/16	百度发布新一代大语言模型、生成式 AI 产品文心一言，文心一言定位于人工智能基座型的赋能平台，将助力金融、能源、媒体、政务等千行百业的智能化变革。
2023/3/16	微软正式宣布推出 Microsoft 365 Copilot，Copilot 用于辅助用户在 Microsoft 365 应用和服务中生成文档、电子邮件、演示文稿等，由 OpenAI 的 GPT-4 技术驱动，像一个助手一样出现在 Microsoft 365 应用的侧边栏，作为一个聊天机器人，让 Office 用户可以随时召唤它，在文档中生成文本、根据 Word 文档创建 PowerPoint 演示文稿，甚

至帮助使用 Excel 中的数据透视表等功能。

2023/3/21 GTC 2023 的主题演讲上, 英伟达的创始人黄仁勋提出如何将人工智能模型应用到芯片设计制造领域, 将硅材料的工艺向着 1nm 的理论极限推进, 刺激算力芯片行业行情。

2023/3/31 意大利个人数据保护局 (Garante) 于当地时间 3 月 31 日宣布, 即日起禁止使用聊天机器人 ChatGPT, 并限制 OpenAI 处理意大利用户信息, AI 安全走强。

2023/4/11 《生成式人工智能服务管理办法(征求意见稿)》发布, 促进生成式人工智能技术健康发展和规范应用

2023/4/28 ChatGPT 在意大利恢复使用, AI 监管较之前有好转

2023/5/10 谷歌发布了全新一代的大语言模型 PaLM2, 与 GPT-4 对抗。据介绍, 他们用了大量数学和科学数据集对 PaLM 2 进行训练, 相比 2022 年的 PaLM 模型, PaLM 2 在多语言处理、推理和编码能力上有了很大的提升。PaLM 2 支持 100 多种不同的语言和 20 多种编程语言, 包括 JavaScript、Python、Prolog、Verilog、Fortran 等, 支持多模态的 PaLM 2 还能看懂和生成音视频内容。

2023/5/25 AI 芯片大厂英伟达 (NVIDIA) 公布了超出市场预期的 2024 财年第一季度财报, 同时还给出了大超市场预期的第二季度预测, 带动英伟达股价盘后暴涨 24.65%, 这也将进一步推动英伟达稳坐全球市值最高的芯片厂商。

2023/6/14 欧洲议会以压倒性多数的投票结果通过欧盟《人工智能法案》草案。该法案对 AI 系统进行风险分类, 限制深度伪造, 并对 ChatGPT 等生成式人工智能提出了更高透明度的要求。

2023/6/20 网信办发布境内深度合成服务算法备案清单, 包括 41 则深度合成服务算法备案信息。体现了我国对此轮人工智能引领的科技变革的重视, 针对生成式算法的法规标准化步伐正在加速推进。

2023/7/19 网信办发布《生成式人工智能服务管理暂行办法》, 预计 2023 年 8 月 15 日开始实行。国内对生成式人工智能服务监管持积极态度, 实行包容审慎和分类分级监管。《办法》中提出鼓励生成式人工智能技术在各行业、各领域的创新应用, 生成积极健康、向上向善的优质内容, 探索优化应用场景, 构建应用生态体系。并鼓励生成式人工智能算法、框架、芯片及配套软件平台等基础技术的自主创新, 平等互利开展国际交流与合作, 参与生成式人工智能相关国际规则制定。

2023/7/19 Meta AI 在周二发布了最新一代开源大模型 Llama 2, 同时宣布与微软云服务 Azure 合作, 向全球开发者首发基于 Llama 2 模型的云服务。

2023/9/25 OpenAI 宣布推出 GPT-4V, GPT-4V 是继 GPT-3 后的最新一代自然语言处理模型, 增加了两项新功能: 语音输入和图像输入, 使 ChatGPT 具备更强大的创造力和语言理解能力

2023/10/17 美国商务部工业和安全局 (BIS) 发布了针对芯片的出口禁令新规, 对于中国半导体的制裁进一步升级, 更加严格地限制了中国购买重要的高端芯片。

2023/10/18 网信中国发布《全球人工智能治理倡议》, 各国应在人工智能治理中加强信息交流和技术合作, 形成具有广泛共识的人工智能治理框架和标准规范, 不断提升人工智能技术的安全性、可靠性、可控性、公平性。

2023/11/1 首届全球人工智能 (AI) 安全峰会召开, 由包括中国在内的与会国共同达成的《布莱切利宣言》正式发表。这是全球第一份针对人工智能这一快速新兴技术的国际性声明, 旨在关注对未来强大人工智能模型构成人类生存威胁的担忧, 以及对人工智能当前增强有害或偏见信息的担忧。

2023/11/7 OpenAI 推出 ChatGPT 自定义功能, 用户无需编码就可为特定目的创建自己版本的 ChatGPT, 以构成不久后发布的 GPT 商店 (GPT Store)。

2023/11/29

生成式 AI 初创公司 Pika Labs 融资筹集 5500 万美元。Pika Labs 宣布推出视频生成器 Pika 1.0。据悉，Pika 1.0 采用了全新的人工智能模型，可以生成和编辑三维动画、动漫、卡通和电影等不同风格的视频。

2023/12/8

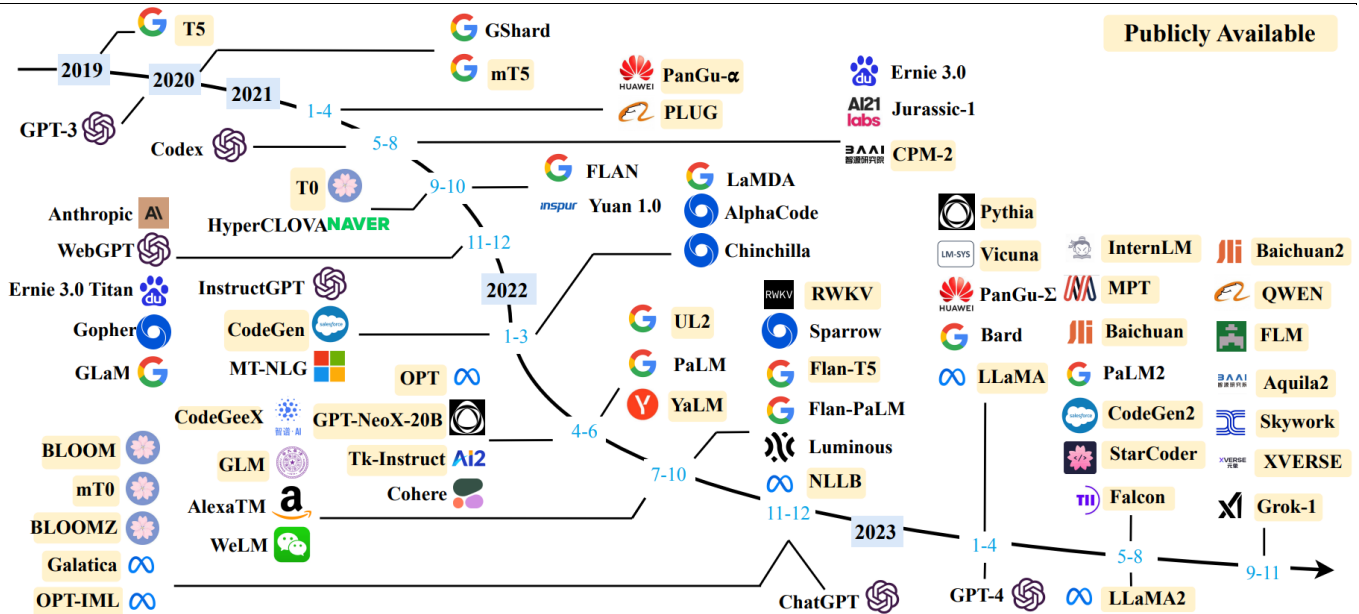
欧洲议会、欧盟成员国和欧盟委员会的谈判代表就欧盟《人工智能法案》达成协议，通过了全球首部人工智能领域的全面监管法规。

资料来源：澎湃新闻，腾讯科技，腾讯网，IT之家，人民网，《生成式人工智能服务管理暂行办法》，财经网，广东省人工智能产业协会，新浪财经，网信中国，观察者网，前瞻网，中国经济网，财联社，界面新闻，第一财经，信达证券研发中心

二、全球大模型风起云涌，中国 AI 行业正在崛起

全球大型语言模型的行业发展进入加速期，2023 年开启大模型之战。从 2020 年 Open AI 的 GPT-3 到 2023 年的 GPT-4，我们看到了该领域创新速度之快。不同公司、科研机构所研发的模型争奇斗艳，表明了这个行业的竞争性和多样性，OpenAI 的 ChatGPT、谷歌的 Gemini 以及百度的 ERNIE 等都是这个领域的重要模型。AI 技术日益大众化使得更广泛的用户能够接触到这些强大便捷的工具，AI 研究与现实世界应用的融合也日益明显，众多公司将这些模型整合到消费者和企业产品中，推动了技术的发展及其实际应用价值。

图 2：全球大模型发展历程



资料来源：Wayne Xin Zhao 等《A Survey of Large Language Models》，信达证券研发中心

海外大模型行业率先起航，OpenAI 引领行业发展。国外大模型起步较早，谷歌于 2017 年研发出 transformer 架构，成为大模型行业的基石。2022 年 11 月，OpenAI 推出的 ChatGPT 更是用户数快速增长，月活迅速破亿，让 AI 大模型走入大众视野，大模型算力和参数量不断迭代升级。目前 Open AI 扮演着开创者的角色，其他大模型厂商快速追赶。自推出 chatGPT 以来，Open AI 不断革新，11 月 7 日，OpenAI 首次开发者大会发布了最新模型 GPT-4 Turbo，可以定制个人版本 ChatGPT 的平台和 GPT Store，开发者还能与 OpenAI 进行收入分成。谷歌作为深度学习的领导者，推出语言模型 PaLM2 和 Gemini 大模型。除 HellaSwag 数据集外，Gemini 其它基准测试领先 GPT-4。开源社区和开源模型也是大模型发展的重要力量。Meta 于 2023 年 7 月发布了免费商用版开源大模型 LLaMA2，同时 Meta 将与微软合作，Llama 2 将在 Azure 和 Windows 上登录。我们认为目前多模态大模型仍是主要发展趋势，后续预计各家厂商将聚焦于 AI 大模型走向规模化应用。

 请阅读最后一页免责声明及信息披露 <http://www.cindasc.com> 7

表 2：国外大模型简介

公司	发布时间	大模型	简介
OpenAI	2023.11	GPT-4Turbo	GPT-4 Turbo 进行了六大升级，分别是更长的上下文长度、更强的控制、模型的知识升级、多模态、模型微调定制和更高的速率限制。定价方面，GPT-4 Turbo 在输入方面比 GPT-4 便宜 3 倍，而输出方面便宜 2 倍。
	2022.04	DALL·E 2	DALL·E2 可以进行修复，在 DALL·E1 的基础上，将文本到图像生成应用在图像更细粒度的级别上。用户可以从现有的图片开始，选择一个区域，让模型对图像进行编辑。另一个功能是生成图像不同变体，用户上传一张图像，然后模型创建出一系列类似的变体。
	2023.11	Whisper3	OpenAI 发布了一系列开源模型——升级版的自动语音识别（ASR）模型——Whisper3。这一模型具备多语言支持，可以将音频内容快速准确地转录成文本，并具备独特的时间戳功能，使其适用于制作字幕等应用。
Google	2023.05	PaLM 2	PaLM 2 在一个涵盖 100 多种语言的语料库上训练，相较此前的模型能够理解更细微的措辞。同时改进了逻辑推理以及对编写和调试代码的支持，在包括 JavaScript、Python、Prolog、Verilog 和 Fortran 等 20 种编程语言上训练。
	2023.12	Gemini	Gemini 1.0 分为 Ultra、Pro 和 Nano 三个版本，具有复杂多模态推理能力，经过训练，可以同时识别并理解文本、图像、音频等，因此它能更好地理解具有细微差别的信息，回答与复杂主题相关的问题，尤其擅长解释数学和物理等复杂科目中的推理。
Meta	2023.07	LLaMa-2	Llama 2 模型系列包含 70 亿、130 亿和 700 亿三种参数变体。相比 Llama 1，Llama 2 的训练数据多了 40%，上下文长度也翻倍，并采用了分组查询注意力机制。具体来说，Llama 2 预训练模型在 2 万亿的 token 上训练，精调 Chat 模型在 100 万人类标记数据上训练。

资料来源：界面新闻，机器之心，站长之家，快科技，信达证券研发中心

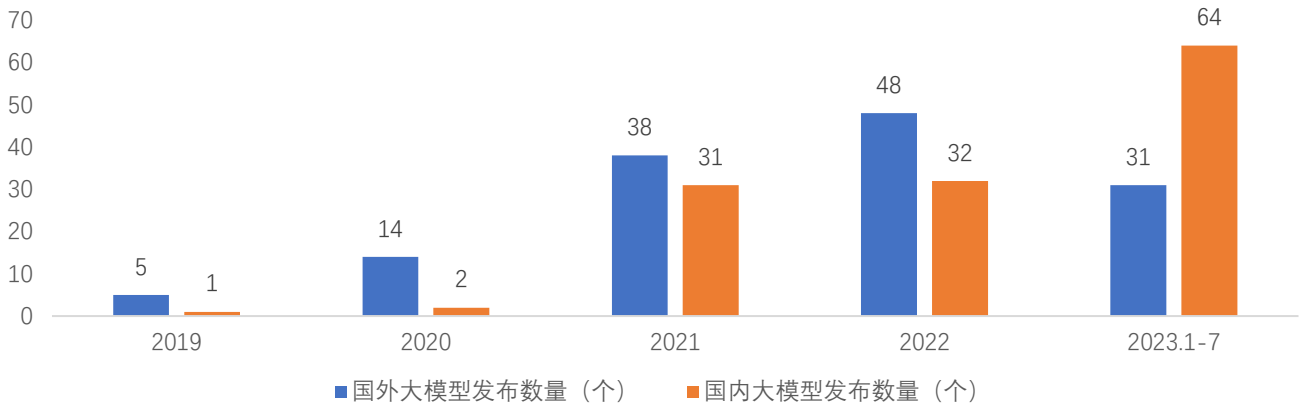
人工智能大模型正掀起新一轮科技发展热潮，重塑生产方式、优化产业结构、提升生产效率的赋能效应日益显现，通用人工智能的序幕已经拉开。随着 ChatGPT 在全球范围内获得显著的成功和关注，中国的主要科技公司（如阿里巴巴、百度、腾讯、华为）、兴起的创业公司（如质谱华章、百川智能、Minimax）、以及在人工智能领域已建立声誉的企业（如科大讯飞、商汤科技），连同学术界和研究机构（如智源研究院、清华复旦等大学）也开始积极投入大模型的开发。根据赛迪顾问数据，截至 2023 年 7 月底，国外大模型发布数量累计达 138 个；23 年 1-7 月国产大模型增长迅速，发布大模型 64 个。当前，国内这些大型模型还处在探索和成长的初级阶段，它们在性能和用户友好度方面还有待市场的进一步验证。我们预计在中国，大规模模型市场的竞争布局要完全成形还需要时间，但考虑到互联网巨头们在人工智能领域的深厚积累，他们可能在这场竞赛中占据先机。

国内产学研用等各方积极布局、大力投身大模型研发应用。百度 CEO 李彦宏指出，截至 2023 年 10 月底国内已经发布 238 个大模型，通过国家网信办备案的大模型超过 20 个，大模型已进入规模落地应用阶段。据赛迪研究院数据，2023 年我国大语言模型市场规模将达到 132.3 亿元，增长率达 110%。在 2023 年 12 月 22 日举行的全国信息技术标准化技术委员会人工智能分委会全体会议上，“大模型标准符合性测试”结果公布，百度文心一言、腾讯混元大模型、360 智脑、阿里云通义千问四家大模型率先通过测试。

从竞争动态和国际趋势，中国的大模型行业很可能会展现出与全球市场相似的模式。首先，具备强大的技术力量、资金支持、人才储备和应用场景的大企业，可能会在通用大模型领域占据主导地位。与此同时，中小企业面临着

自己的挑战和机遇，他们可以依靠自身在特定细分市场和数据处理方面的优势，成为专注于垂直特定领域的关键参与者。另外，由于模型训练和用户调用对计算资源的庞大需求，考虑到成本效益和资源可获取性，这些公司可能会更倾向于与云服务提供商建立合作关系。

图 3：国内外大模型发布情况



资料来源：赛迪顾问，信达证券研发中心

表 3：国内大模型简介

公司	发布时间	大模型	简介
百度	2023.10	文心一言 4.0	作为全球大厂中首个发布的生成式 AI 产品，百度“文心一言”于今年 3 月 16 日开启邀测，10 月，百度发布文心一言最新迭代产品一文心一言 4.0。11 月 1 日，百度上线文心一言专业版文心大模型 4.0 是迄今为止最强大的文心大模型实现了基础模型的全面升级，文心大模型 4.0 的理解和生成能力的提升幅度相近，而逻辑和记忆能力的提升则更大，逻辑的提升幅度达到理解的近 3 倍，记忆的提升幅度也达到了理解的 2 倍多。以理解能力为例，文心大模型 4.0 已具备前后乱序、模糊意图、潜台词等复杂提示词的理解力，还能胜任更专业的代码理解力与调试任务。
阿里云	2023.10	通义千问	继 4 月正式发布通义千问大模型后，10 月 31 日，阿里发布了通义千问 2.0 版本。此外，通义千问 APP 在各大手机应用市场正式上线，所有人都可通过 APP 直接体验最新模型能力。通义千问 2.0 在 10 个权威测评中全面超越 GPT-3.5 和 Llama2，加速追赶 GPT-4。通义千问 2.0 在复杂指令理解、文学创作、通用数学、知识记忆、幻觉抵御等能力上均有显著提升。目前通义千问的综合性能已经超过 GPT-3.5，加速追赶 GPT-4 与此同时，基于通义大模型训练的 8 大行业模型组上线，他们分别是通义灵码-智能编码助手、通义智文-AI 阅读助手、通义听悟-工作学习 AI 助手、通义星尘-个性化角色创作平台、通义点金-智能投研助手、通义晓蜜-智能客服、通义仁心-个人专属健康助手、通义法睿-AI 法律顾问。
华为	2023.07	盘古大模型 3.0	盘古大模型强调在细分场景的落地应用，聚焦产品研发、软件工程、生产供应等五大价值场景，致力于深耕行业包括政务、金融、制造、煤矿、铁路、制药、气象等领域。盘古大模型 3.0 是个面向行业的大模型系列，包括“5+N+X”三层架构。其中，LO 层包括自然语言、视觉、多模态、预测、科学计算五个基础大模型，提供满足行业场景中的多种技能需求。L1 层是 N 个行业大模型，华为云既可以提供使用行业公开数据训练的行业通用大模型，包括政务、金融、制造、矿山、气象等大模型；也可以基于行业客户的自有数据，在盘古大模型的 LO 和 L 层上，为客户训练自己的专有大模型。L2 层为客户提供了更多细化场景的模型，更加专注于政务热线、网点助手、先导药物筛选、传送带异物检测、台风路径预测等具体行业应用或特定业务场景，为客户提供“开箱即用”的模型服务。
腾信	2023.09	混元	腾讯混元大语言模型在 2023 腾讯全球数字生态大会上正式亮相，并宣布通过腾讯云对外开放，这是腾讯首次披露的通用大语言模型。混元大模型拥有超千亿参数规模，预训练语料超 2 万亿 tokens，具有强大的中文理解与创作能力逻辑推理能力，以及可靠的任务执行能力。在多个场景下，腾讯混元大模型已经能够处理超长文本，通过位置编码优化技术，混元大模型对于长文处理效果和性能得到了提升。并且混元大模型还具有识别“陷阱”的能力，简单来说就是通过强化学习方法拒绝被“诱导”。目前，基于混元大模型的多模态对话 Bot 产品“腾讯混元助手”微信小程序现已开放内测申请，用户可申请排队体验，审核通过将短信提醒。小程序的简介显示，“腾讯混元助手”的功能包括 AI 问答、AI 绘画，可以回答各类问题，也能处理多种任务，如获取

			知识、解决数学问题、翻译、提供旅游攻略工作建议等。
京东	2023.07	言犀	<p>京东言犀大模型是立足于产业研发的，它融合了70%通用数据与30%数智供应链原生数据，具有“更高产业属性、更强泛化能力更多安全保障”的优势，致力于各种产业场景，解决真实的产业问题。言犀大模型已经在消费导购、商家经营、客服售后、医疗问诊等多个供应链场景中试点接入。此外，在内部经营管理方面，京东尝试将大模型应用于系统代码辅助编写实现了20%以上的效率提升；也开始测试AIGC自动生成商品营销图文的能力目前已推广至2000多个零售三级品类。</p>
抖音集团	2023.08	云雀	<p>云雀大模型的定位是一个多媒体内容生成和内容理解的AI大模型，该模型基于字节神经网络加速器开发，可以根据用户输入的图片、视频、音频、文字等信息，生成各种类型和风格的多媒体内容，如视频剪辑、音乐配乐、滤镜特效、字幕翻译等。它还可以对输入的多媒体内容进行分析和评价，如标签、分类、质量、风格等。</p>
百川智能	2023.04	Baichuan 系列	<p>百川智能成立于2023年，由搜狗创始人王小川创立。经过半年多时间发展百川智能便接连发布Baichuan-7B/13B, Baichuan2-7B/13B 四款开源可免费商用大模型及Baichuan-53B、Baichuan2-53B 两款闭源大模型，平均每28天就会发布一款新的大模型。百川智能透露，Baichuan-7B/13B 两款开源大模型在多个权威评测榜单均名列前茅，累计下载量超过六百万次。Baichuan2-13B 在 MMLU、CMMLU、MedQA、USMLE 等几大权威评估基准中领先 LLaMA2.10月30日，百川智能宣布推出Baichuan2-192K大模型，其上下文窗口长度高达192K，能够处理约35万个汉字。Baichuan2-192K 将以API调用和私有化部署的方式提供给企业用户。目前百川智能已经启动该大模型的API内测，并开放给法律、媒体、金融等行业的核心合作伙伴。</p>
网易有道	2023.07	子曰	<p>网易有道在今年7月首次推出了国内首个专注于教育领域的垂直大模型“子曰”，并发布了六个基于该大模型的创新应用，包括“LLM 翻译”、“虚人“语法精讲”、“AI Box”和“文档问答”语教练”、“AI 作文指导”、相比于通用大模型，教育垂类大模型[子曰]拥有更专业的预训练语料，可以依据用户在学习场景下的需求，向其提供对话式大语言模型，帮助用户答疑解惑基于该模型，有道为不同学习场景设计了定制化的模型，以实现模型与场景的高度契合</p>
昆仑万维	2023.08	天工 Skywork	<p>天工由昆仑万维与国内领先的AI团队奇点智源联合研发，是国内首个对标ChatGPT的双千亿级大语言模型，也是一个AI搜索引擎，一个对话式AI助手天工大模型用过通过自然语言与用户进行问答式交互，AI生成能力可满足文案创作、知识问答、代码编程、逻辑推演、数理推算等多元化需求。天工大模型拥有强大的自然语言处理和智能交互能力，能够实现个性化AI搜索、智能问答、聊天互动、文本生成、编写代码、语言翻译等多种应用场景并且具有丰富的知识储备，涵盖科学、技术、文化、艺术、历史等领域。</p>
知乎	2023.04	知海图 AI	<p>4月，知乎推出首个中文大模型“知海图AI”，该大模型在成熟大模型CPM-Bee基础上研发，拥有千亿级参数，具备更强的逻辑推理能力，更快的训练和推理速度。“知海图AI”在公司业务提效上收益显著，其在分层、分类、兴趣理解、搜索等业务场景广泛应用大模型能力，人工标注量降低了90%以上业务准召效果普遍提升了15%以上。除了优秀的大语言模型能力，“知海图AI”还具备强大的多模态处理能力，助力知乎实现会员故事智能配图，赋能会员业务发展。值得一提的是，知乎教育业务也开展了基于大模型的产品创新，通过AI智能申论批改、智能作业批改、AI教务领航员，极大提升了教学效率和教学质量。</p>
科大讯飞	2023.10	讯飞星火 3.0	<p>2023年以来，科大讯飞已经完成了对星火大模型的多次升级。6月9日，星火大模型V1.5发布，开放式问答、多轮对话及数学能力升级；8月15日，星火大模型升级至V2.0，突破代码能力、多模态交互再升级。10月24日，科大讯飞发布星火大模型V3.0，实现七大能力持续提升。具体来看，星火V3.0进一步升级了数学自动提炼规律、小样本学习、代码项目级理解能力、多模态指令跟随与细节表达等能力，进一步提升星火的落地应用能力。此外，讯飞星火大模型V3.0新增了AI人设功能，它可以根据性格模拟、情绪理解、表达风格来形成一个初始人设，在结合特定知识学习、对话记忆学习，形成一个更个性化的AI人设。</p>
智谱华章	2023.08	智谱清言	<p>智谱清言是基于智谱AI自主研发的中英双语对话模型ChatGLM2用户可通过苹果商店App Store、安卓主流应用市场(华为、OPPO、vivo、应用宝及小米)下载或在微信小程序搜索“智谱清言”进行体验。智谱清言可作为用户的智能助手，在工作、学习和日常生活中为用户解答各类问题，完成各种任务。具备的丰富能力包括通用问答、媒体写作、学习、职场和编程等。例如在媒体写作方面，智谱清言可以帮助用户完成文章创作、新闻选题、微博宣发产品PR、APP种草、短报编写、公众号推文等任务。</p>

商汤科技	2023.04	日日新 SenseNova 大模型	日日新大模型涵盖语言大模型和多模态全栈技术能力。商汤基于 1.5 万亿 token 和 6000 亿中文高质量数据，依托拥有 30000+张 GPU 的商汤 AI 大装置 SenseCore，训练 20/100/120/180B 不同场景下大模型，推出国内首个综合能力超越 GPT-3.5 的大语言模型以及文生图大模型，服务上千家客户并实现模型相关业务规模 6 倍增长。日日新大模型拥有提供自然语言、内容生成、自动化数据标注、自定义模型训练等多种能力，其中还包括商量 SenseChat、秒画 SenseMirage、如影 SenseAvatar、琼宇 SenseSpace、格物 SenseThings、明眸等功能强大且易用的内容生成平台产品。
	2023.09、2023.11	智脑、奇元	今年 11 月，360 智脑官方宣布，360 大模型“奇元大模型”通过备案落地从大模型定位和应用角度来看，奇元大模型具备充足的灵活性和可扩展性，商业化和产品定位以 B 端用户为主，后期将会聚焦更多的商业化应用和垂直领域帮助用户提升工作效率。据了解，“360 智脑大模型”9 月已获批面向公众开放至此，360 成为国内首家两个大模型均通过备案的科技企业。目前，360 智脑大模型已全面接入 360 搜索、360 浏览器等互联网全端应用。360 智脑大模型集成了 360GPT 大模型、360 多模态大模型技术能力，具备生成创作、多轮对话逻辑推理等十大核心能力、数百项细分功能，覆盖大模型全部应用场景。
复旦大学	2023.04	MOSS	MOSS 是国内第一个插件版对话语言模型，能使用搜索引擎、图像生成模型、方程求解器等外部工具，为用户提供越来越多的服务。整体而言，MOSS 基于公开的中英文数据训练，目前已经拥有 200 亿参数，具有和人类对话的能力，并可以通过与人类交互实现迭代优化。但 MOSS 在部署和训练上投入的资本还非常有限，相较于千亿级的 ChatGPT，大概只是其规模的十分之一，所以 MOSS 还有很多事实类知识记不住，思维链能力也相对较差。

资料来源：数据观，腾讯网，中国网科学，人民网，AIBetas，澎湃新闻，IT 之家，中国网，商汤科技，OSCHINA，信达证券研发中心

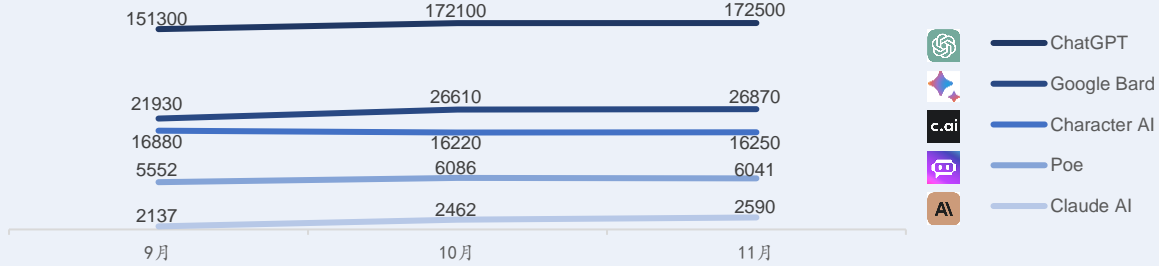
三、生成式 AI 应用及展望

AI 大模型高速迭代，多模态 AI 应用持续探索迭代。相比于纯文本大语言模型，多模态模型也有自己的基座模型、预训练模型以及带有涌现能力的模型，如在基础模型 CLIP、DALIE 的基础上也已衍生出类似 Flamingo、PaLM-E 等视觉多模态语言模型，通用大语言模型在高速迭代中。从 AI 应用类型来看，目前大语言模型应用不只局限于 NLP 文本生成，多模态大模型已经可以处理来自不同模态（如视频、图像、语音、文本等）的多种信息，如自然语言处理、多模态推理、计算机视觉和音频处理等，实现模态之间的转换和统一。各行业也正期待着在多模态大模型方向上能有类似“ChatGPT 时刻”的出现。

表 4：不同应用场景 AI 模型访问量排名

趋势	访问量排名	产品名称	产品简介	23.09-11 访问总量	访问量月增长量
AI 聊天机器人					
AI 聊天机器人用户数量庞大且趋于平稳，ChatGPT 作为先行者占据绝大部分市场份额。	1	ChatGPT	一个基于 OpenAI 的 GPT 模型的交互式聊天机器人，提供富有洞察力的文本回答和生成。	49.59 亿	0.22%
	2	Google Bard	谷歌的创新语言 AI 模型，旨在提供新颖的信息探索和创意内容创作服务。	7.54 亿	1.00%
	3	Character AI	提供个性化聊天体验，允许用户与定制的 AI 角色进行互动对话。	4.88 亿	-0.03%
	4	Poe	由 OpenAI 支持的 AI 平台，集成多个模型，旨在提供灵活的问答和学习工具。	1.77 亿	-0.75%
	5	Claude AI	Anthropic 的 AI 对话系统，强调对话的理解、推理能力和提高可解释性。	0.72 亿	5.20%

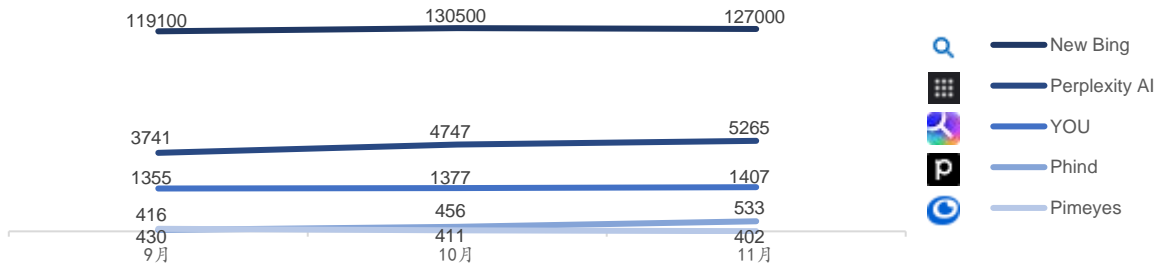
2023年9-11月AI聊天机器人网站访问量（万）


AI 搜索引擎

排名	搜索引擎	描述	访问量 (亿)	环比变化 (%)
1	New Bing	微软推出的新一代搜索引擎, 结合传统搜索功能和 ChatGPT 式的交互式聊天能力。	37.67	-2.71%
2	Perplexity AI	专注于提供深度、解释性强的问答服务, 利用先进的语言模型来理解和回答复杂问题。	1.38	10.92%
3	YOU	一种以隐私保护为特色的搜索引擎, 注重用户个性化和数据保护。	0.41	2.16%
4	Phind	使用图像识别技术来提供视觉搜索服务, 帮助用户识别和发现物品和地点。	0.14	17.08%
5	Pimeyes	一个面部识别搜索引擎, 用于在网络上查找人脸图像, 强调版权保护和个人隐私。	0.12	-2.18%

AI 搜索引擎呈多元化和专业化的趋势: 大平 New Bing 仍占主导地位, 新兴和专业的服务商快速增长。

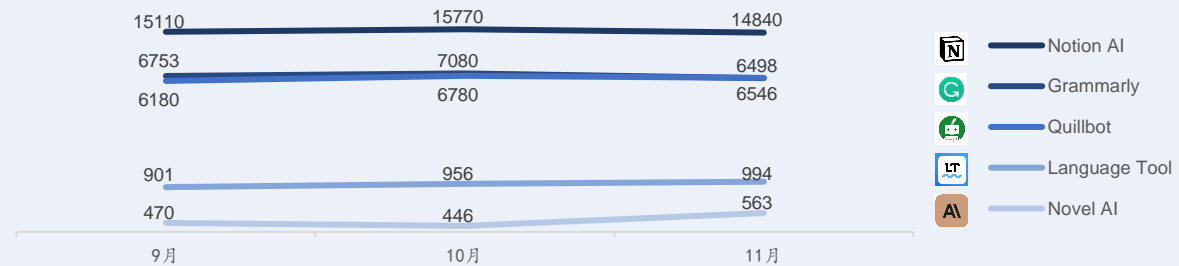
2023年9-11月AI搜索引擎网站访问量(万)


AI 文字编辑

排名	文字编辑工具	描述	访问量 (亿)	环比变化 (%)
1	Notion AI	Notion 的 AI 功能, 提供智能笔记整理、内容生成和工作流自动化。	4.57	-5.88%
2	Grammarly	一个文本校正工具, 专注于提高写作质量, 包括语法检查和风格改进。	2.03	-8.22%
3	Quillbot	提供改写和摘要生成工具, 帮助用户改善和简化他们的写作。	1.95	-3.45%
4	Language Tool	一个多语言文本校正工具, 提供语法、拼写和风格建议。	0.29	4.00%
5	Novel AI	利用人工智能技术来辅助创意写作, 尤其是小说创作。	0.15	26.14%

纠错、整理、摘要等传统文字编辑辅助工具小幅下滑, 文字创作平台快速增长。

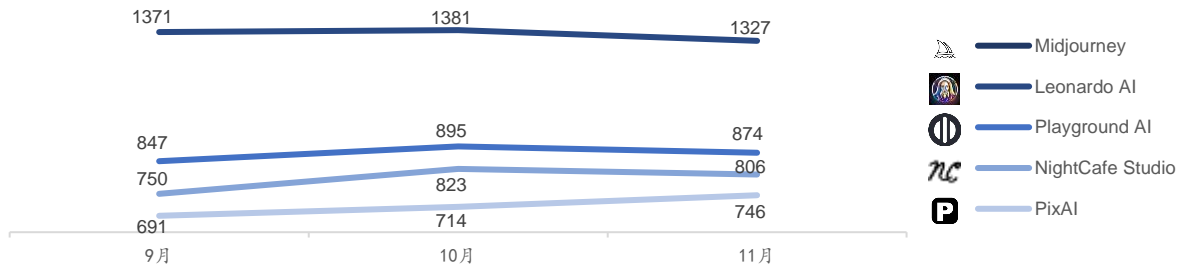
2023年9-11月AI文字编辑网站访问量(万)


AI 图片创作

综合类图片生成工具小幅下滑，专业化、个性化图片生成工具增长。

排名	名称	描述	市值	涨跌幅
1	Midjourney	一个创新的 AI 绘图工具，可以根据文本提示生成逼真的图片。	0.51 亿	-8.60%
2	Leonardo AI	使用预先训练的人工智能模型生成独特的艺术作品或训练用户自己的模型。	0.41 亿	-3.92%
3	Playground AI	一个在线 AI 图像创造工具，允许用户通过上传图像和使用简单的文本提示来生成 AI 图像。	0.26 亿	-2.35%
4	NightCafe Studio	提供了 AI 艺术生成器，使用户能够使用人工智能快速创造艺术作品。	0.24 亿	-2.12%
5	PixAI	一个以二次元风格为主的 AI 绘图工具。	0.22 亿	4.48%

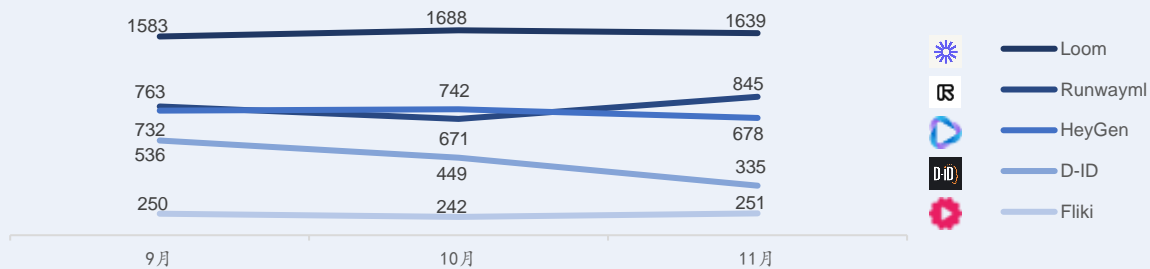
2023年9-11月AI图片创作网站访问量(万)



视频创作大类功能日渐丰富，市场竞争较为激烈；文生视频产品呈增长趋势。

排名	名称	描述	市值	涨跌幅
1	Loom	利用 AI 录制视频并与用户的团队成员和客户分享，以提高工作效率	0.49 亿	-2.95%
2	Runwayml	提供包括文本到图像、图像到图像、视频帧插值和自定义 AI 训练等工具。	0.23 亿	25.91%
3	HeyGen	一个 AI 视频生成工具，能够轻松制作高质量的视频，利用 AI 生成的虚拟角色和语音。	0.22 亿	-8.59%
4	D-ID	一个专注于 AI 生成视频创作平台，能够制作逼真的对话 AI 代理，旨在人性化地与数字界面进行交互。	0.13 亿	-25.42%
5	Fliki	一个 AI 视频和语音内容创作平台，提供了文本到视频和文本到语音的转换功能。	0.07 亿	3.73%

2023年9-11月AI视频创作网站访问量(万)

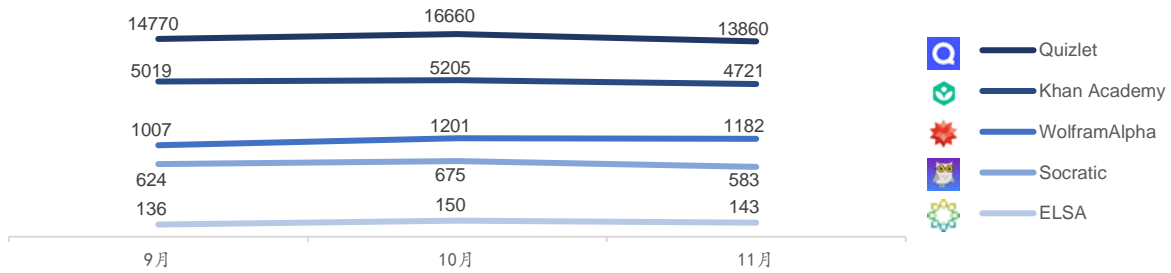


AI 教育学习

AI 教育学习平台热度呈下滑趋势。

排名	名称	描述	市值	涨跌幅
1	Quizlet	一个内置了基于 ChatGPT 的人工智能导师 Q-Chat 的互动对话型学习平台。	4.53 亿	-16.81%
2	Khan Academy	一个内置了一位随时可用的导师和助教 Khanmigo 的学习平台。	1.49 亿	-9.30%
3	Wolfram Alpha	一个使用 Wolfram 的突破性算法、知识库和 AI 技术得出出专家级答案的答案引擎。	0.34 亿	-1.58%
4	Socratic	一个由谷歌人工智能提供支持适用于所有科目的学习平台。	0.19 亿	-13.59%
5	ELSA	ELSA，英语语言语音助手，是一款有趣且引人入胜的应用，专为帮助提高英语发音而设计。	0.04 亿	-4.28%

2023年9-11月AI教育学习网站访问量(万)

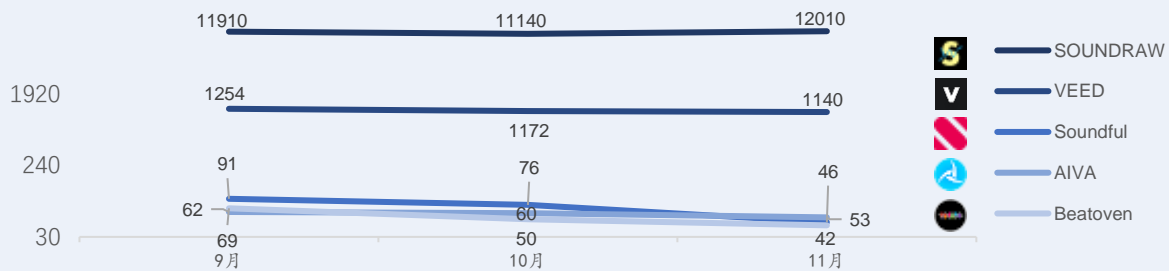


SOUNDDRAW
持续领跑市场，小份额平台市场占比持续下滑。

AI 音乐创作

排名	平台	描述	估值	变化率
1	SOUNDDRAW	一款综合型人工智能音乐生成器，可帮助创作者为其项目制作定制音乐。	35.08 亿	7.81%
2	VEED	一款视频编辑工具，内置 AI 音乐生成功能，可通过文本和曲风免费生成音乐。	0.36 亿	-2.73%
3	Soundful	一个人工智能音乐生成器平台，使内容创作者和音乐艺术家能够创建曲目并通过音乐获利。	0.02 亿	-39.86%
4	AIVA	一款 AI 音乐生成助手，可生成 250 多种不同风格的新歌曲。	0.02 亿	-12.09%
5	Beatoven	一个使用先进的 AI 音乐技术来生成独特的基于情绪的音乐，以适应视频的每个部分的平台。	0.02 亿	-15.73%

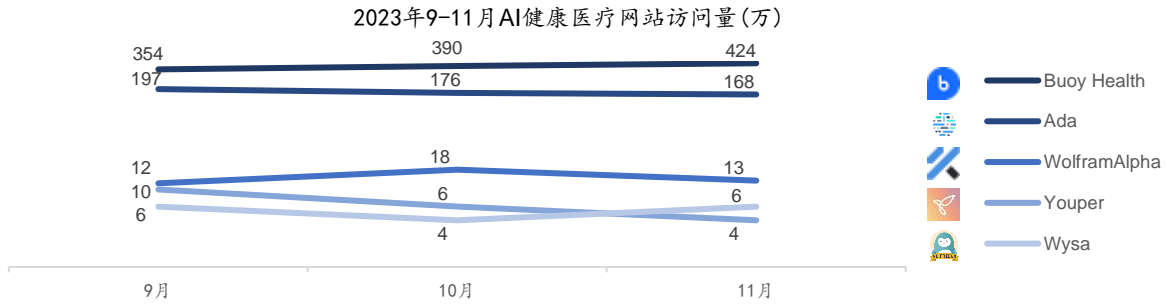
2023年9-11月AI音乐创作网站访问量(万)



AI 健康医疗

AI 健康医疗用户存量较少，在医疗诊断和心理健康子类中，Buoy 和 Wysa 较竞争对手 Ada 和 Youper 更具优势。

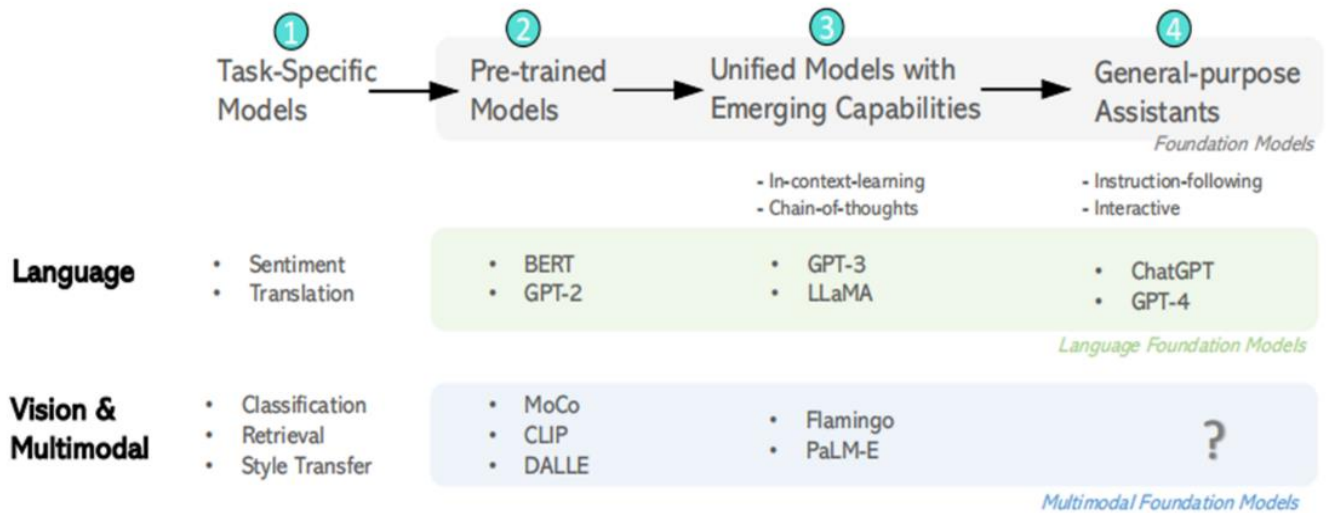
排名	平台	描述	估值	变化率
1	Buoy Health	Buoy Health 是一个能够检查用户症状并明确护理选择，指导用户走向健康的人工智能平台。	0.11 亿	8.72%
2	Ada	Ada 是一款通过智能技术支持更好的健康结果和卓越的临床表现的人工智能。	0.05 亿	-4.55%
3	FitnessAI	FitnessAI 是一款人工智能驱动的健身训练应用程序，它能根据用户的个人目标、健身水平和健康背景为其生成个性化训练方案。	43.19 万	-31.42%
4	Youper	Youper 是另一款人工智能驱动的心理应用，它依赖于认知行为疗法 (CBT) 的实践和方法。	19.55 万	-35.75%
5	Wysa	Wysa 是一款移动应用程序，它采用多种方法提供心理健康护理，同时使用人工智能和真人为用户提供所需的护理。	16.07 万	55.24%



资料来源: similarweb, 信达证券研发中心

生成式 AI 在文本和图像生成领域持续演进。文本、图像生成领域一直是生成式 AI 发展的重要方向，在扩散模型成为目前主流之前，GAN（Generative Adversarial Networks）技术一直是比较流行的生成模型之一，英伟达和微软在早期均有所探索，如英伟达的 StyleGAN 系列。后面随着技术的迭代发展，OpenAI 陆续推出了 DALL-E 的三个版本，也诞生了 Stable Diffusion 系列、Midjourney 等爆款文生图领域相关产品，而 DALL-E3 目前的生成效果较好，其原生构建在 ChatGPT 基础上，用 ChatGPT 来构建、拓展和优化 Prompt，在生成画面的丰富度、细节、清晰度等方面表现较为优秀，人们对文图大模型领域的探索一直在进行。

图 4：文本和视觉多模态大模型的发展轨迹对比



资料来源: Chunyaan Li 等《Multimodal Foundation Models: From Specialists to General-Purpose Assistants》，信达证券研发中心

文生图应用迭代效果优秀，DALL·E-3 和 Midjourney V6 的图像细节和光影效果突出。2023 年 9 月，OpenAI 推出最新得文生图模型 DALL·E-3，在语义理解、颗粒处理、图像质量等方面，比之前发布的 DALL·E-2 实现大幅度提升，可生成写实、二次元、平面、创意、朋克、3D 等多种类型，图片质量、丰富度和准确度相较竞争对手某些维度更优，在语意、细节理解上表现更好。除了较好得生图效果之外，DALL·E 3 的最大特点是与 ChatGPT 集成，它原生构建在 ChatGPT 之上，用 ChatGPT 来创建、拓展和优化 prompt。能更好地理解上下文，并且处理较长的 prompt 效果会更好，将 prompt 中的措辞都尽可能表现出来。同时在安全性方面，OpenAI 在 DALL·E3 上投入了大量工作，包括制定强有力的安全措施，以防止创建「有害」的图像。

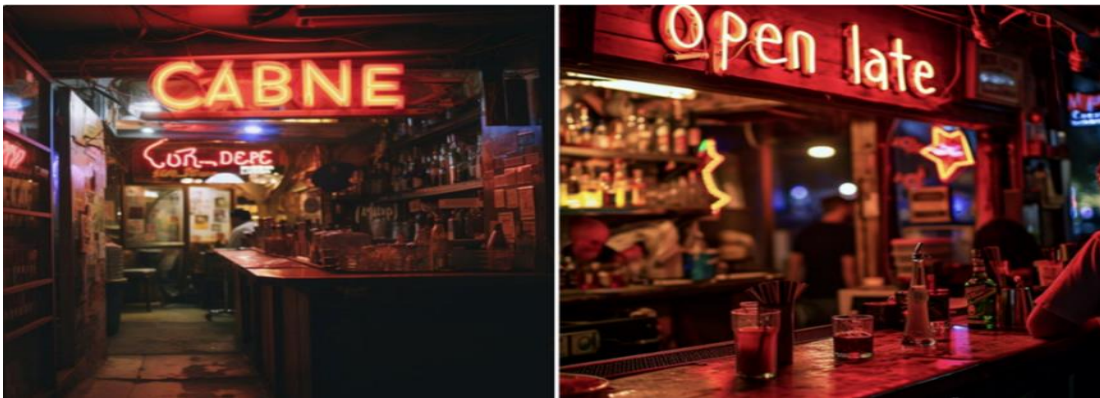
图 5: DALL·E3 较 DALL·E2 在图片丰富度、细节方面的进步对比



资料来源: OpenAI 官网, 信达证券研发中心

另一款海外较流行的文生图应用 Midjourney 已于近期更新了 V6 版本, 更新后的版本图像更加真实、细腻、光影细节明显。总结来说, V6 共有 5 大升级: 更精确且更长的提示响应、改进了连贯性和模型知识、图像生成和混合 (remix) 得到了优化、新增了基础文字绘制功、upscale (放大器) 功能得到增强, 具有 'subtle' 和 'creative' 两种模式, 分辨率提升两倍。

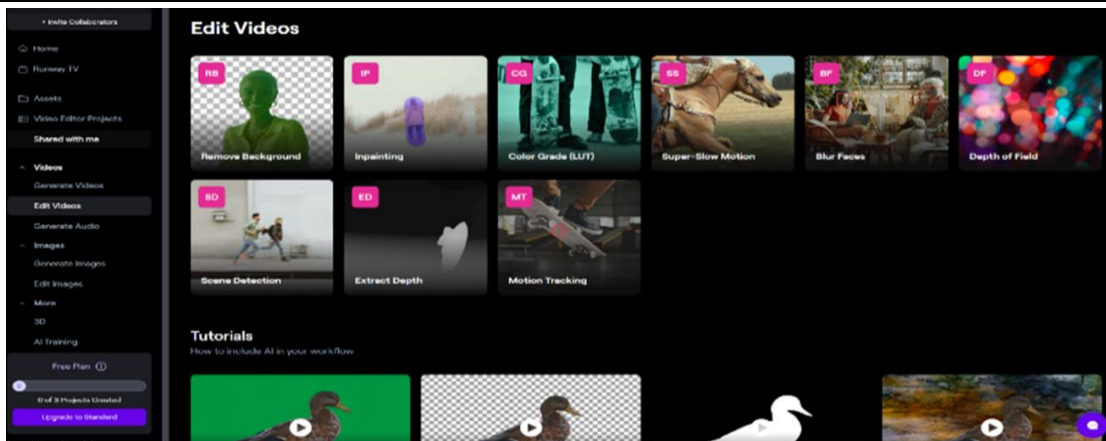
图 6: Midjourney V5.2、V6 生成图像细节对比



资料来源: 新智元公众号, 信达证券研发中心

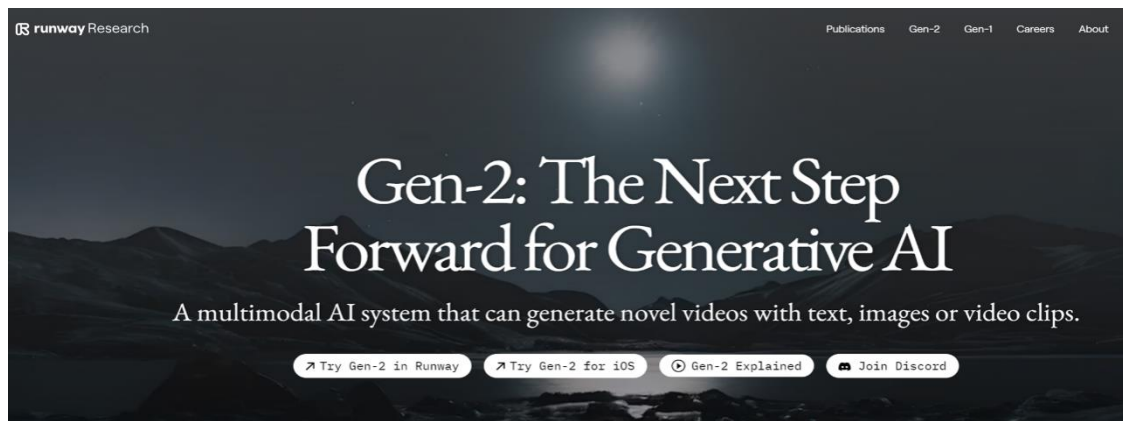
文生视频应用迭代进展超预期, Runway 和 Pika 产品引领行业革新。在文生视频领域的进展远比想象中迅速: Runway 先后发布 Gen1、Gen2, 推出的 Motion Brush 则在可靠性上向前了一步, Stability AI 也在近期发布了首个 Text-to-Video 模型 Stable Video Diffusion。

图 7: Runway 视频剪辑工具功能



资料来源: Runway 官网, 信达证券研发中心

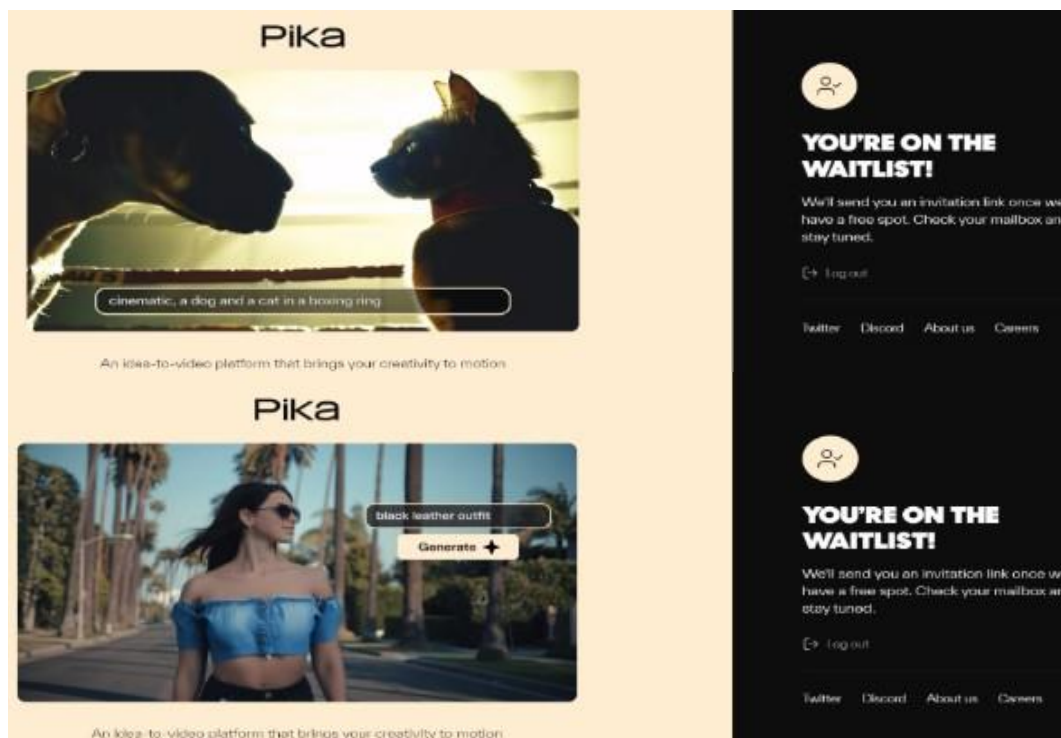
图 8: Runway Gen-2 模型



资料来源: Runway 官网, 信达证券研发中心

2023 年 7 月, Pika Labs 在 Discord 推出服务器, 并在几个月时间内收获了 50 万用户, 逐渐成为可以同 Runway 同台竞技的黑马公司。11 月底, Pika Labs 正式推出了最新一代视频生成模型 Pika 1.0。Pika 1.0 不仅对视频生成质量进行了提升, 还支持 3D 动画、动漫或电影等多种风格视频。Pika Labs 目前已经累计完成 3 轮、共 5500 万美元融资, 最新估值为 2.5 亿美元。Nat Friedman、Daniel Gross 和 Lightspeed 分别领投了 Pika 的 Pre-Seed、Seed 和 A 轮融资, 此外, 众多硅谷明星投资人、AI 领域重要公司 CEO 或核心成员也参与了投资。Pika 1.0 亮点功能具体包括: 1) 文本生成视频/图像生成视频; 2) 面向上传视频进行风格转换; 3) 扩展视频画布; 4) 原有视频内容通过 AI 更改: 如更换服装颜色、添加角色、添加道具等; 5) 延长视频: 使用 AI 扩展现有视频剪辑的长度几秒钟, 但延长视频的一致性目前存在问题待改进, 且随着延长时间的增加, 目前总体视频效果有所降低。

图 9: Pika 文生视频内容



资料来源: Pika 官网, 信达证券研发中心

文生 3D 应用产业化难度略高, 仍在探索阶段。RODIN 模型是由微软研究院提出的一种文字生成 3D 模型, 首次

实现了在 3D 训练数据上利用生成扩散模型自动生成 3D 数字化身；其核心功能在于，只需一张图片甚至一句文字描述，即可秒级生成 3D 化身。尽管目前 RODIN 仍属于试验性产品，生成结果仅限于半身人像，微软表示未来将不断扩展技术能力和适用场景。相较于过去的 3D 生成应用，RODIN 模型融合了 NeRF 及英伟达 EG3D 等工作，能够生成更丰富的 3D 细节，提高了精细度。

在游戏行业，降本增效是关键，而低成本可定制化的 3D 建模技术或成为一个热点。RODIN 模型的核心能力在于输入文字或图片，即可直接生成 3D 的人物模型，并且支持后续修改装扮和形象，契合游戏行业对人物模型的需求。相较于过去的 3D 生成应用，这种技术具有更高的精细度，能够提供更个性化的定制化体验。在 RODIN 被提出后，2023 年 5 月，OpenAI 也发布了文字生成 3D 模型 Shap.E，并在 Github 上开源；这一模型已经能够生成 3D 资产，为文本生成 3D 领域的发展增添了新的动力。

图 10：图/文生成 3D 人物模型过程示意




资料来源：微软亚洲研究院《3D 生成扩散模型 RODIN，秒级定制 3D 数字化身》，信达证券研发中心

后续展望：多模态应用能力圈有望不断拓展。GPT-4V 的输入支持纯文本、单个图像-文本对、交错图像-文本对三种情况，遵循文字说明、视觉指向和视觉参考提示、视觉+文本等多种提示技术，可以逐渐达到多种能力，例如：识别不同种类的东西、理解图像中人与物体之间的空间关系、确定位置、计算数量；以及对场景文本、表格图表的推理能力、多语言文本识别描述、代码能力、多模态尝试理解和推理能力等。

在与多模态系统的人机交互中，指向特定空间位置是一项基本能力，例如进行基于视觉的对话，GPT-4V 可以很好地理解直接画在图像上的视觉指针。同时，微软研究人员也探索了 GPT-4V 在时间和视频理解上的应用，探索时序预测、时序排序、时序定位、时序推理和基础时序理解等能力有助于衡量模型在一系列静态图像中理解事件顺序、预测未来事件发生和分析随时间变化的活动的的能力。

图 11：GPT-4V 对指向性提示以及逻辑思考能力


Prompt:
How much did I pay for tax?
Where is the information?



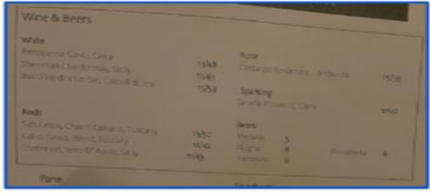
Prompt Continued

GPT-4V:
In the first receipt, you paid \$3.72 for tax. The information is located at the bottom of the receipt. In the second receipt, you paid \$42.23 for tax. The information is located near the bottom of the receipt, under the "Sales Tax" line. In the third receipt, you paid \$4.58 for tax. The information is located at the bottom of the receipt, under the "TAX" line. In total, you paid \$50.53 in tax across all three receipts.

Prompt:
How much should I pay for the beer on the table according to the price on the menu?



Here is a photo of the menu.




GPT-4V:
According to the menu, the price for a Magna beer is \$6. Since there are two Magna beers on the table, the total cost for the beers would be \$12.

资料来源: Zhengyuan Yang 等《The Dawn of LMMs: Preliminary Explorations with GPT-4V》, 信达证券研发中心


图 12: GPT-4V 对视频理解和时间序列的理解能力

Prompt:
Predict what will happen next based on the images.




GPT-4V:
Based on the images, it appears that the player in the white jersey is taking a penalty kick. The next likely event is that the player will kick the ball towards the goal, attempting to score. The goalkeeper in the black jersey will try to block the shot.

Prompt:
Analyze the images and decide what the next steps should be.




GPT-4V:
The images show the process of making sushi rolls. The first image shows rice being mixed with vinegar in a bowl. The second image shows the rice being spread onto a sheet of seaweed. The third image shows the addition of salmon and cucumber to the roll. The next steps would be to roll the sushi using a bamboo mat, cut the roll into smaller pieces, and serve with soy sauce and wasabi.

Prompt:
For opening a door, decide the correct order in which they should be displayed.



GPT-4V:
The correct order would be:
1. a - Person approaching the door
2. b - Person opening the door
3. c - Person entering through the door.

Prompt:
For closing a door, decide the correct order in which they should be displayed.



GPT-4V:
Correct order:
1) c - Person standing in doorway
2) b - Person starting to close door
3) a - Person closing door completely

资料来源: Zhengyuan Yang 等《The Dawn of LMMs: Preliminary Explorations with GPT-4V》, 信达证券研发中心

四、投资建议

受益于全球生成式 AI 持续发展, 文生图/视频/音乐/代码等多方面应用的迭代更新有望长期给诸多行业带来降本增效。建议关注, AI 大模型端: 腾讯、百度、阿里巴巴、昆仑万维、科大讯飞等;

应用端: 1) AI+办公: 微软、万兴科技、福昕软件、金山办公;

2) AI+游戏: 紫天科技 (AI 游戏先行者, 率先发力互动剧)、昆仑万维 (新推出 AI Agent 开发平台, 天工大模型六大矩阵看点丰富)、巨人网络 (《原始征途》表现优异, 休闲品类有望受益新渠道)、掌趣科技 (与悠米、

行者 AI、layabox 等合作布局 AI+游戏领域)、神州泰岳(游戏出海表现稳健,非游戏板块拥抱 AI)、盛天网络(游戏业务稳定基本盘,带带电竞及给麦带来增量贡献)、恺英网络、三七互娱、吉比特、宝通科技、名臣健康、姚记科技、顺网科技、汤姆猫、星辉娱乐等;

3) **AI+影视**: 上海电影(集团支持+IP 优势)、华策影视(影视剧储备丰富+微短剧布局)、光线传媒(动画电影受益于文生视频进展)等;

4) **AI+出版**: 果麦文化、南方传媒、中南传媒、凤凰传媒、皖新传媒等;

5) **AI+教育**: 世纪天鸿等;

6) **AI+营销**: 三人行、芒果超媒、蓝色光标、分众传媒、易点天下、兆讯传媒、浙文互联、遥望科技等。

7) **IP 资源**: 奥飞娱乐、中文在线;

8) **美股标的**: 微软、谷歌、Adobe、Unity 等;

9) **港股标的**: 腾讯、美图公司、拼多多、阜博集团、柠萌影视、小米集团、小鹏汽车等。

五、风险因素

1. 生成式 AI 公司研发及资本开支不及预期

生成式 AI 行业在经历了短期热潮后,热门应用的流量及活跃用户数或有所下滑,后续或需要更多有实质性提升工作效率和娱乐体验等方面的应用落地,用户持续性付费或可成为 AI 公司持续投入资本开支的关键;

2. 生成式 AI 应用迭代不及预期

目前在多模态应用中,文生图片、文生视频等均已出现了一系列应用,生成的效果目前仍无法直接应用于大部分行业中,产品仍在持续迭代优化,后续存在迭代不及预期风险。

研究团队简介

冯翠婷，信达证券传媒互联网及海外首席分析师，北京大学管理学硕士，香港大学金融学硕士，中山大学管理学学士。2016-2021年任职于天风证券，覆盖互联网、游戏、广告、电商等多个板块，及元宇宙、体育二级市场研究先行者（首篇报告作者），所在团队曾获21年东方财富Choice金牌分析师第一、Wind金牌分析师第三、水晶球奖第六、金麒麟第七，20年Wind金牌分析师第一、第一财经第一、金麒麟新锐第三。

凤超，信达证券传媒互联网及海外团队高级研究员，本科和研究生分别毕业于清华大学和法国马赛大学，曾在腾讯担任研发工程师，后任职于知名私募机构，担任互联网行业分析师。目前主要负责海外互联网行业的研究，拥有5年的行研经验，对港美股市场和互联网行业有长期的跟踪覆盖。主要关注电商、游戏、本地生活、短视频等领域。

刘旺，信达证券传媒互联网及海外团队高级研究员。北京大学金融学硕士，北京邮电大学计算机硕士，北京邮电大学计算机学士，曾任职于腾讯，一级市场从业3年，创业5年（人工智能、虚拟数字人等），拥有人工智能、虚拟数字人、互联网等领域的产业经历。

李依韩，信达证券传媒互联网及海外团队研究员。中国农业大学金融硕士，2022年加入信达证券研发中心，覆盖互联网板块。曾任职于华创证券，所在团队曾入围2021年新财富传播与文化类最佳分析师评比，2021年21世纪金牌分析师第四名，2021年金麒麟奖第五名，2021年水晶球评比入围。

白云汉，信达证券传媒互联网及海外团队成员。美国康涅狄格大学金融学硕士，曾任职于腾讯系创业公司投资部，一级市场从业2年。后任职于私募基金担任研究员，二级市场从业3年，覆盖传媒互联网赛道，拥有游戏领域产业链资源。2023年加入信达证券研究所，目前主要专注于微软、网易等美股研究以及结合海外映射对A股港股的覆盖。

分析师声明

负责本报告全部或部分内容的每一位分析师在此申明，本人具有证券投资咨询执业资格，并在中国证券业协会注册登记为证券分析师，以勤勉的职业态度，独立、客观地出具本报告；本报告所表述的所有观点准确反映了分析师本人的研究观点；本人薪酬的任何组成部分不曾与，不与，也将不会与本报告中的具体分析意见或观点直接或间接相关。

免责声明

信达证券股份有限公司(以下简称“信达证券”)具有中国证监会批复的证券投资咨询业务资格。本报告由信达证券制作并发布。

本报告是针对与信达证券签署服务协议的签约客户的专属研究产品，为该类客户进行投资决策时提供辅助和参考，双方对权利与义务均有严格约定。本报告仅提供给上述特定客户，并不面向公众发布。信达证券不会因接收人收到本报告而视其为本公司的当然客户。客户应当认识到有关本报告的电话、短信、邮件提示仅为研究观点的简要沟通，对本报告的参考使用须以本报告的完整版本为准。

本报告是基于信达证券认为可靠的已公开信息编制，但信达证券不保证所载信息的准确性和完整性。本报告所载的意见、评估及预测仅为本报告最初出具日的观点和判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会出现不同程度的波动，涉及证券或投资标的的历史表现不应作为日后表现的保证。在不同时期，或因使用不同假设和标准，采用不同观点和分析方法，致使信达证券发出与本报告所载意见、评估及预测不一致的研究报告，对此信达证券可不发出特别通知。

在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议，也没有考虑到客户特殊的投资目标、财务状况或需求。客户应考虑本报告中的任何意见或建议是否符合其特定状况，若有必要应寻求专家意见。本报告所载的资料、工具、意见及推测仅供参考，并非作为或被视为出售或购买证券或其他投资标的的邀请或向人做出邀请。

在法律允许的情况下，信达证券或其关联机构可能会持有报告中涉及的公司所发行的证券并进行交易，并可能会为这些公司正在提供或争取提供投资银行业务服务。

本报告版权仅为信达证券所有。未经信达证券书面同意，任何机构和个人不得以任何形式翻版、复制、发布、转发或引用本报告的任何部分。若信达证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，信达证券对此等行为不承担任何责任。本报告同时不构成信达证券向发送本报告的机构之客户提供的投资建议。

如未经信达证券授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。信达证券将保留随时追究其法律责任的权利。

评级说明

投资建议的比较标准	股票投资评级	行业投资评级
本报告采用的基准指数：沪深300指数（以下简称基准）； 时间段：报告发布之日起6个月内。	买入 ：股价相对强于基准20%以上；	看好 ：行业指数超越基准；
	增持 ：股价相对强于基准5%~20%；	中性 ：行业指数与基准基本持平；
	持有 ：股价相对基准波动在±5%之间；	看淡 ：行业指数弱于基准。
	卖出 ：股价相对弱于基准5%以下。	

风险提示

证券市场是一个风险无时不在的市场。投资者在进行证券交易时存在赢利的可能，也存在亏损的风险。建议投资者应当充分深入地了解证券市场蕴含的各项风险并谨慎行事。

本报告中所述证券不一定能在所有的国家和地区向所有类型的投资者销售，投资者应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专业顾问的意见。在任何情况下，信达证券不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任，投资者需自行承担风险。