

2023年03月26日

GPT 模型走向平台化，「手转 AI」大幕拉开

AGI（通用人工智能）专题之三

► ChatGPT 着手打造内部生态，“模型即平台”时代开启

2022年3月，OpenAI 开始逐步开放内部插件，并允许第三方在 GPT 模型内创建插件。我们认为这实际上是划时代的举措：

1) 开放内部插件的意义，可以理解为用插件拓展 GPT 模型自身能力的边界，为模型添加“眼睛和耳朵”。尽管 GPT4 的语言模型被微软称作“早期 AGI”，但依然存在数据库时间滞后、知识幻觉（输出缺乏依据的内容）等 LLM 常见的问题。通过允许第三方插件的接入，大模型得以联网，用户可以在 ChatGPT 内部借助插件完成更多样的任务。

2) 一旦形成第三方插件生态，GPT 模型将成为 AI 时代的操作系统本身，可类比移动互联网时代诞生的操作系统 IOS 及安卓。不同之处在于，GPT 模型并不是依靠硬件终端变化来实现代际变革，而是依靠人工智能技术为用户提供的实际便利场景去完成用户习惯的迁移。如果说 ChatGPT 模型的伟大之处是开创性地发现了 RLHF+transformer+精细标注数据这一菜谱，通过工程化落地让人们广泛地接触到 AI（类比智能手机出现），那么推出内生插件则是更有野心、也更能看到划时代和商业意义的一步：一个基于 AI 的全新系统生态近在眼前。

► 外部成熟应用将向 AI 平台迁移，开启“手转 AI”浪潮

AI 加持下，GPT 内部插件带来的用户体验优于直接使用外部的原生应用，用户的使用场景得以迁移到 GPT 模型内部。目前插件商店内部的第一批应用主要由 OpenAI 邀请的第三方开发，包括电商、办公协作、旅游、在线餐厅预订平台、计算软件等。随着外部成熟应用的交互圈层扩大，GPT 与用户日常生活将结合得更加紧密，参考移动互联网时代的端转手趋势，主流应用即将展开向 AI 平台的迁移，即迅速适应并布局 GPT 内的 AI 插件，探索 AI 加持下自身应用的新场景。这实际上是新技术与流量的互惠：成熟应用依赖 GPT 的 AI 技术提高用户体验，而 GPT 则借助这些应用吸引更多用户，网络效应进一步加强了这种相互促进的趋势，从而提升 GPT 平台的生态价值。

► 插件开发门槛大幅降低，行业格局可能重塑

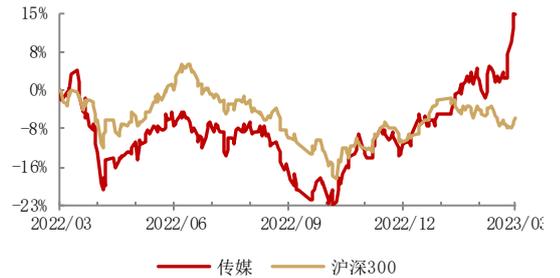
AI 技术的加持下，用户开发 GPT 插件几乎不需要技术门槛。开发者只需简单描述对 API 的需求，模型会完成代码编写、身份验证、数据处理等工作。这意味着 UGC（用户生成内容）回潮，玩法、用户体验和差异化成为下一阶段的关键：

1) 由于用户和开发者不必担心复杂的技术实现细节，每个人都可以根据自己的需求创建插件，活跃的社区氛围下用户粘性将提升，插件生态将更加繁荣。2) 对于开发者而言，创意、

评级及分析师信息

行业评级：推荐

行业走势图



分析师：赵琳

邮箱：zhaolin@hx168.com.cn

SAC NO: S1120520040003

玩法、差异化体验、与实际场景的融合度等成为了决定自身插件能否取得先发优势的核心因素。我们认为未来文娱、电商、教育等多个行业格局可能重塑。

投资建议：

1) 随着海外大模型率先平台化、生态化，国产大模型的复现需求变得紧迫。模型层受益标的包括百度、华为（产业链）、阿里、三六零等明确具备大模型落地意愿及能力的厂商，同时追踪腾讯等厂商的相关技术进展。

2) 应用层：境外业务占主导、海外用户基础大、有实际应用场景的公司会尽快接入 GPT 生态，先发优势下，我们认为这类厂商的估值将直接受益于 OpenAI 的波浪式创新。此外，海外映射逻辑下，关注布局游戏、浏览搜索、电商、本地生活等业务，并已有 AIGC 相关技术实践的标的。

风险提示

国产大模型落地效果不及预期风险； AI 产品道德及监管风险。

正文目录

1. ChatGPT 推出内部插件，“模型即平台”时代开启.....	4
1.1. 开放三大基础插件，GPT4 开始联网.....	5
1.2. 第三方插件：开启成熟应用向 AI 平台迁移的序幕.....	9
1.3. 未来的插件生态：开发门槛极大降低，UGC 属性强.....	10
2. 从“端转手”到“手转 AI”，国产替代需求更加紧迫.....	11
3. 投资建议.....	15
4. 风险提示.....	15

图表目录

图 1 “插件商店”出现在 ChatGPT 的新版本中.....	4
图 2 数理插件 Wolfram 加持下的对话页面.....	4
图 3 浏览插件下用户可以就最新的资讯提问.....	5
图 4 利用浏览插件提问奥斯卡获奖演员/影片.....	6
图 5 提问奥斯卡获奖演员/影片并要求赋诗一首.....	6
图 6 代码解释器从图像中提取文字.....	7
图 7 代码解释器将图像转换为 ASCII 图像.....	7
图 8 代码解释器完成机器学习任务（预测下一个词）.....	7
图 9 代码解释器能够识别图像中的人脸.....	7
图 10 代码解释器可以生成音频文件.....	8
图 11 代码解释器生成二维码.....	8
图 12 代码解释器生成行星轨迹（使用 Matplotlib 库）.....	8
图 13 代码解释器能够利用简单的像素画图.....	8
图 14 用户提问气候数据时，数据插件调取联合国数据库并生成表格.....	9
图 15 商店目前已有的插件主要为电商、出行、订餐、资讯等.....	9
图 16 GPT 能够调用多个插件在一段输出内同时解决用户的生活化问题.....	10
图 17 自制一个“代办清单”插件流程只需要三分钟，代码生成、身份验证、交互等工作由 GPT 完成.....	11
图 18 ChatGPT 的技术突破点在于引入了 RLHF（基于人类反馈的强化学习）.....	12
图 19 “文心一言”生成图片.....	13
图 20 “文心一言”生成视频.....	13
图 21 盘古 NLP 大模型.....	14
图 22 盘古 CV 大模型.....	14
图 23 达摩院“文本生成视频大模型”已经上线.....	14
表 1 Github 社区中主流 AI 框架情况（2022.1）.....	12

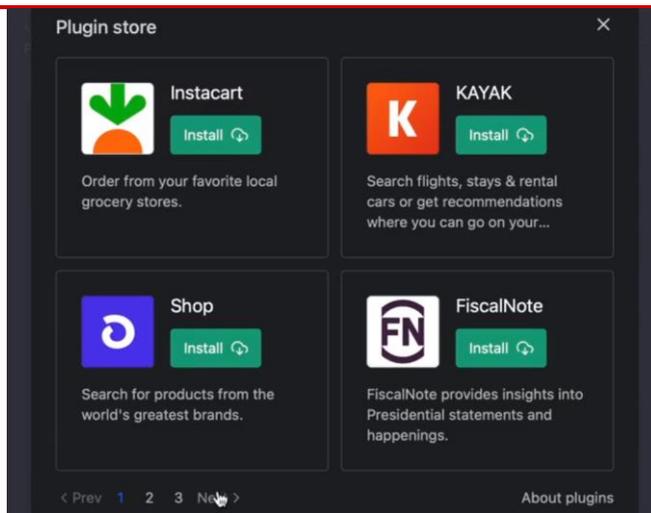
1.ChatGPT 推出内部插件，开启“模型即平台”时代

2022 年 3 月，OpenAI 开始逐步开放内部插件，并允许第三方在 GPT 模型内创建插件。我们认为这实际上是划时代的举措：

1) ChatGPT 开放内部插件的意义，可以理解为用插件拓展 GPT 模型自身能力的边界，为模型添加“眼睛和耳朵”。尽管 GPT4 的语言模型被微软称作“早期 AGI”，但用户目前使用的基于 GPT4 的 ChatGPT 付费版依然只能进行文本的处理，且依然存在数据库时间滞后、知识幻觉（输出缺乏依据的内容）等 LLM 常见的问题。通过允许第三方插件的接入，大模型得以联网，用户可以在 ChatGPT 内部借助插件完成更多样的任务，如资料查找、数据分析、行程预定、发布社交内容等。

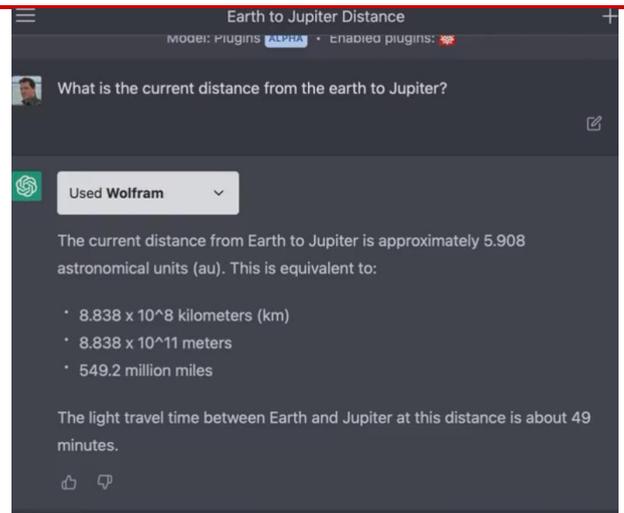
2) 一旦形成第三方插件生态，GPT 模型将成为 AI 时代的操作系统本身。我们可以将其视作围绕 AI 展开的操作系统或底层平台，类比移动互联网时代诞生的操作系统 IOS 及安卓。不同之处在于，GPT 模型并不是依靠硬件终端变化来实现代际变革，而是依靠人工智能技术为用户提供的实际便利场景去完成用户习惯的迁移。因此，开放内部插件最大的意义在于，随着插件与日常生活应用的紧密结合，GPT 模型将不再是少部分人（尽管日活已经突破一亿，ChatGPT 并未对全部国家及地区开放，使用门槛相对较高）使用的对话模型，而是一个所有寻求使用 AI 帮助的用户都必须使用的操作系统，AI 时代的“IOS”诞生。

图 1 “插件商店”出现在 ChatGPT 的新版本中



资料来源：公司官网，华西证券研究所

图 2 数理插件 Wolfram 加持下的对话页面



资料来源：公司官网，华西证券研究所

不同于向外出售 API 接口，内部插件的使用场景仅在对话页面内，实际上将插件应用带来的流量圈存在了 GPT 平台内部。此前 ChatGPT 的变现模式集中于向外出售 API 接口或微调模型，是对外赋能开发者或者 Office 等成熟应用；插件系统推出之后，依靠 GPT 的技术支持，开发者能够以极低的技术成本编写插件，优质插件能够进一步激活用户基础、提升粘性，第三方（如已经成熟的外部应用 expedia）插件更能进一步引流外部用户至 GPT，生态系统已经基本形成。

演示页面中已经出现了插件商店，商业变现潜力极大。目前插件功能还只对申请队列中的开发者和付费用户开放，但我们认为随着功能逐步开放给普通用户，插件商店完全可以复刻现有应用商店模式，优质插件可以按月订阅或买断，外部成熟

APP 会迁移至 GPT 模型内部形成 AI 加持下的插件（如同移动互联网时代的“端转手”）。

我们认为，如果说 ChatGPT 模型的伟大之处是开创性地发现了 RLHF+transformer+精细标注数据这一菜谱，通过工程化落地让人们广泛地接触到 AI（类比智能手机出现），GPT4 模型在多模态能力上推进了 AGI 领域的进展，那么推出内生插件则是更有野心、也更能看到划时代和商业意义的一步：一个基于 AI 的全新系统生态近在眼前。

1.1. 开放三大基础插件，GPT4 开始联网

尽管 GPT4（的语言模型版本）已经被微软认为达到了早期 AGI（通用人工智能）的阶段，模型表现依然受到一定限制。2022 年 3 月对外发布的 GPT-4 在抽象、理解、视觉、编程、数学、医学、法律、理解人类动机和情感等领域表现出色，微软提出它可以被合理地视为 AGI（通用人工智能）系统的早期版本。AGI 意味着智能不仅限于特定领域或任务，而且涵盖了广泛的认知技能和能力。但它们仍然存在一些问题，如缺乏元认知、认知谬误和非理性、训练数据过时等。此外，GPT4 在 ChatGPT 中唯一可以开箱即用的能力就是输出文本，但仍需要用户对提示词进行较多处理。

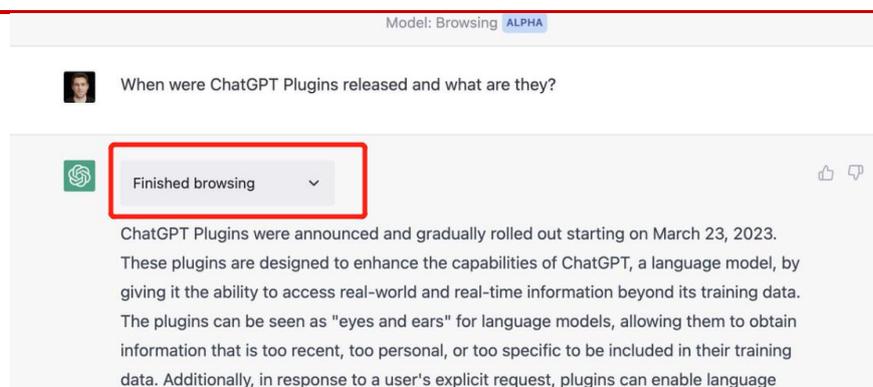
插件的引入可类比为帮助模型补上了“眼睛和耳朵”。2022 年 3 月，OpenAI 开始逐步开放插件，并为开发人员创建插件提供支持。插件不仅允许机器人浏览网页，还可以与特定网站进行交互。我们梳理 OpenAI 提供三个基础插件如下：

1.1.1. 浏览插件：GPT 联网，用户体验实现极大跃迁

浏览插件即 GPT 模型调用 Bing 搜索的 API，从网络检索内容，以此为依据实时推理和输出文本。这意味着大语言模型能够从互联网上读取信息，此前一直被人诟病的封闭性、滞后性得到解决（GPT4 的训练数据库停留在 2021 年）。海量的必应搜索结果扩大了可以讨论的内容范围，超越了训练数据集的时间限制，融入了当下的新鲜信息，一定程度上缓解了大模型常见的“幻觉”问题：大模型常对自己不掌握的知识做出看似合理的荒谬解答。

不同于 NewBing 在搜索页面调用 GPT4，浏览插件是在 ChatGPT 的对话场景内调用 Bing 搜索，实现了反客为主。下图为例，在对 GPT 模型问出涉及最新资讯的问题时，通过调用浏览插件，大模型完成了流畅准确的归总和输出。

图 3 浏览插件下用户可以就最新的资讯提问

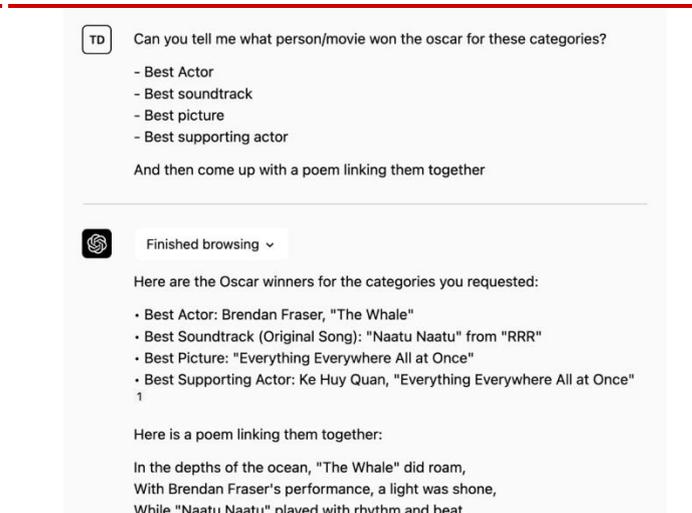
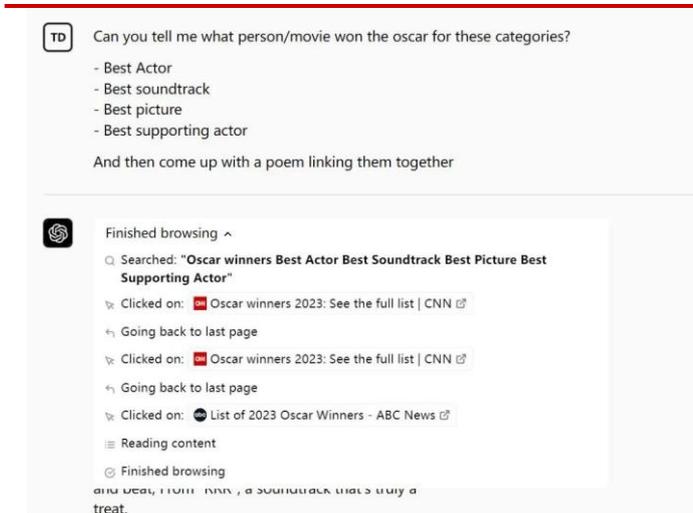


资料来源：ChatGPT，华西证券研究所

从官方示例来看，在进行奥斯卡相关提问时，ChatGPT 检索了最新一届的奥斯卡奖信息，并在信息流中附加链接。若在此基础上要求 GPT 将奖项信息以诗歌方式呈现，并不会影响输出的真实性。

图 4 利用浏览插件提问奥斯卡得奖演员/影片

图 5 提问奥斯卡得奖演员/影片并要求赋诗一首



资料来源：ChatGPT，华西证券研究所

资料来源：ChatGPT，华西证券研究所

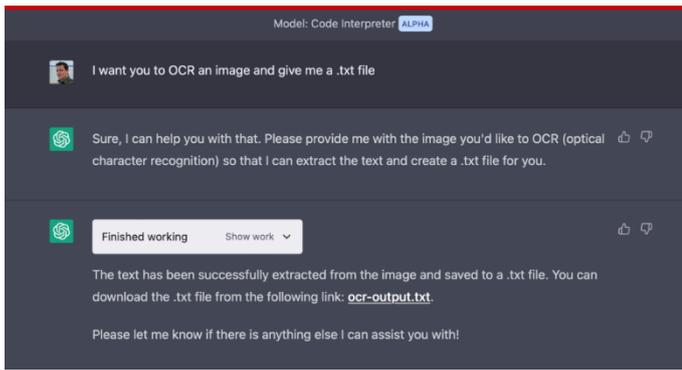
联网的语言模型此前已有 WebGPT、GopherCite、BlenderBot2、LaMDA2 等，但表现并不理想，我们判断主要系抓取网络内容进行输出会影响回答的质量。浏览插件为此在来源可靠性、信息真实性以及防止检索有问题内容的「安全模式」方面做了大量工作，插件在一个独立的服务中运行，与基础设施的其余部分分开。

训练数据集的版权界限依然模糊，但抓取网站数据会事先征求网站许可，法律风险相比封闭训练更低。为了尊重内容创作者并遵守网络规范，ChatGPT 浏览器插件的用户代理 token 是 ChatGPT-User，并配置为尊重网站的 robots.txt 文件。若网站拒绝抓取，GPT 不会将网站内容输出，而不是以任何自动方式抓取网络。OpenAI 还发布了 IP 出口范围，并实施了限速措施以避免向网站发送过多流量。

1.1.2. 代码解释器：GPT 将成为开发平台

代码解释器插件能够帮助用户在 ChatGPT 界面中完成完整的开发工作。用户能够在 GPT 的帮助下通过描述需求，直接生成代码、运行代码、上传和下载从 CSV 数据到图像的文件，并评估输出。此前用户即使可以通过 ChatGPT 创建代码，也需要将其放入另一个环境中进行测试。在新插件的加持下，用户可以在 ChatGPT 中进行大量开发，而无需离开 UI。

图 6 代码解释器从图像中提取文字



资料来源：andrewmayneblog，华西证券研究所

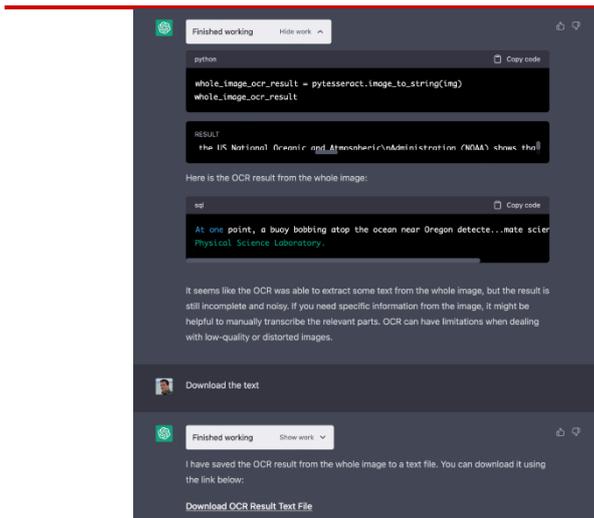
图 7 代码解释器将图像转换为 ASCII 图像



资料来源：andrewmayneblog，华西证券研究所

代码解释器展现了极强的多模态能力，极大降低了用户的开发门槛。根据 OpenAI 提供的信息，代码解释器能够解决定量和定性的数学问题、进行数据分析和可视化、在格式之间转换文件等。我们参考网络用户反馈，除了生成代码之外，代码解释器还可以分析输出，并将其用于另一个函数。这意味着用户可以将不同的代码段串在一起，获取一个代码的输出并将其提供给另一个。从现有尝试来看，用户通过简单易懂的语言描述自己的需求，GPT 可以将其快速完成，目前已完成的任务包括音频及图像生成、画图、识别人脸、自建机器学习预测算法、绘制行星轨迹图、编写小游戏等。

图 8 代码解释器完成机器学习任务（预测下一个词）



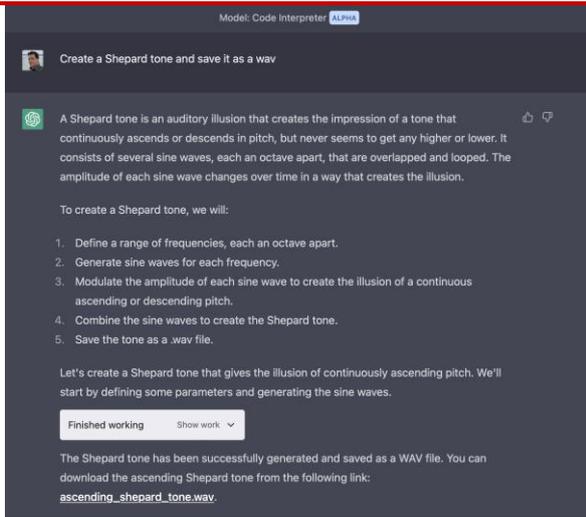
资料来源：andrewmayneblog，华西证券研究所

图 9 代码解释器能够识别图像中的人脸



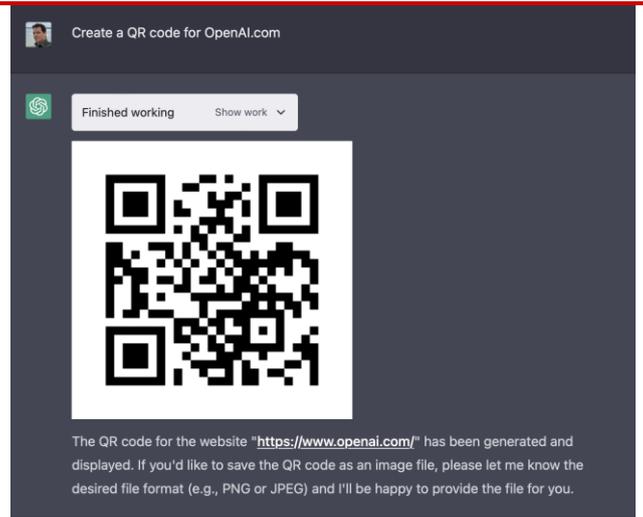
资料来源：andrewmayneblog，华西证券研究所

图 10 代码解释器可以生成音频文件



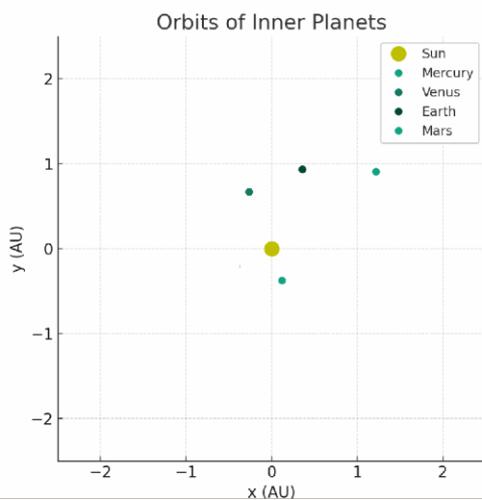
资料来源: andrewmayneblog, 华西证券研究所

图 11 代码解释器生成二维码



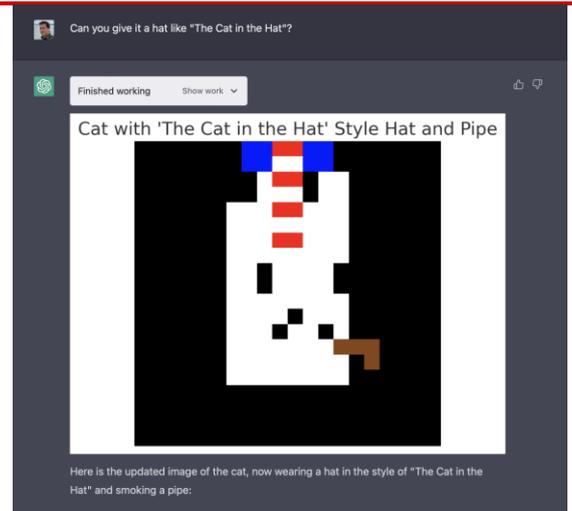
资料来源: andrewmayneblog, 华西证券研究所

图 12 代码解释器生成行星轨迹 (使用 Matplotlib 库)



资料来源: andrewmayneblog, 华西证券研究所

图 13 代码解释器能够利用简单的像素画图

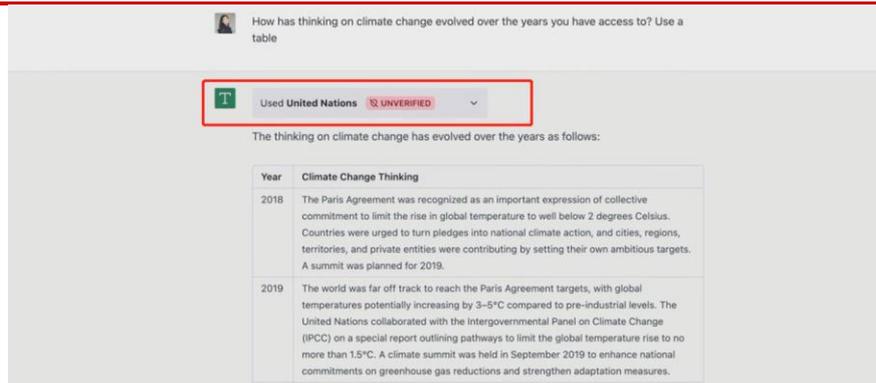


资料来源: andrewmayneblog, 华西证券研究所

1.1.3. 数据插件：模型得以访问开源数据并输出分析

数据插件使 ChatGPT 能够访问开源数据库，并生成分析结果，目前已经向公众开放。用户能够用自然语言的方式表达需求，获取最相关的文档片段，例如文件、笔记、电子邮件或公共文档，并进行数据分析。用户可以自己选取向量数据库（如 Milvus、Pinecone、Qdrant、Redis、Weaviate 或 Zilliz）来索引和搜索文档，添加到他们有权使用的检索插件中，插件在一个独立的服务中运行，与基础设施的其余部分分开，一定程度上规避了数据授权和隐私风险。

图 14 用户提问气候数据时，数据插件调取联合国数据库并生成表格

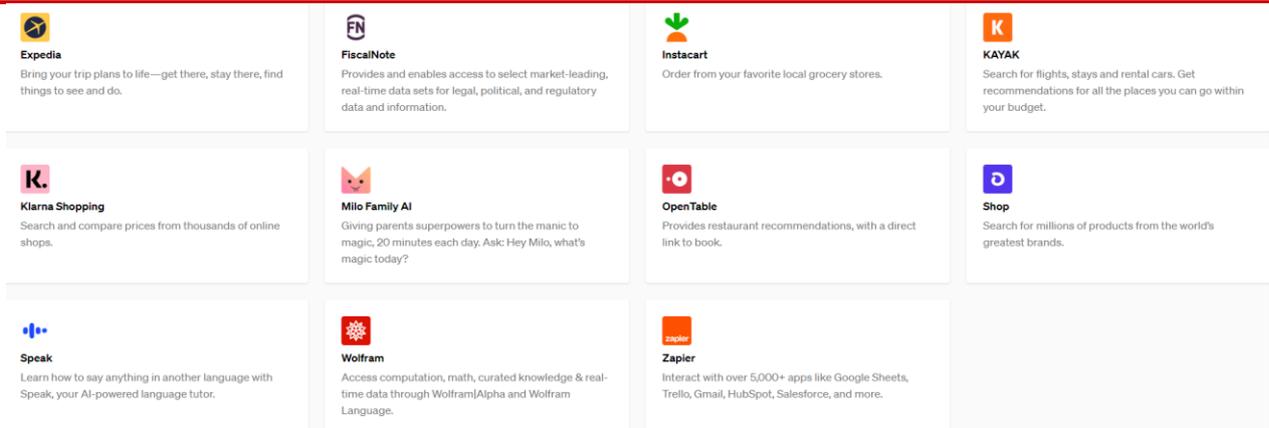


资料来源：ChatGPT，华西证券研究所

1.2. 第三方插件：开启成熟应用向 AI 平台迁移的序幕

GPT 提供的技术支持，让插件带来的用户体验优于直接使用外部的原生应用，用户的使用场景得以迁移到 GPT 模型内部。目前插件商店内部的第一批应用主要由 OpenAI 邀请的第三方开发，包括电商平台 Shopify、办公协作软件 Slack、旅游平台 Expedia、生鲜电商 Instacart、旅游搜索引擎服务商 Kayak、在线餐厅预订平台 OpenTable、自动化平台 Zapier、计算软件 Wolfram 等。

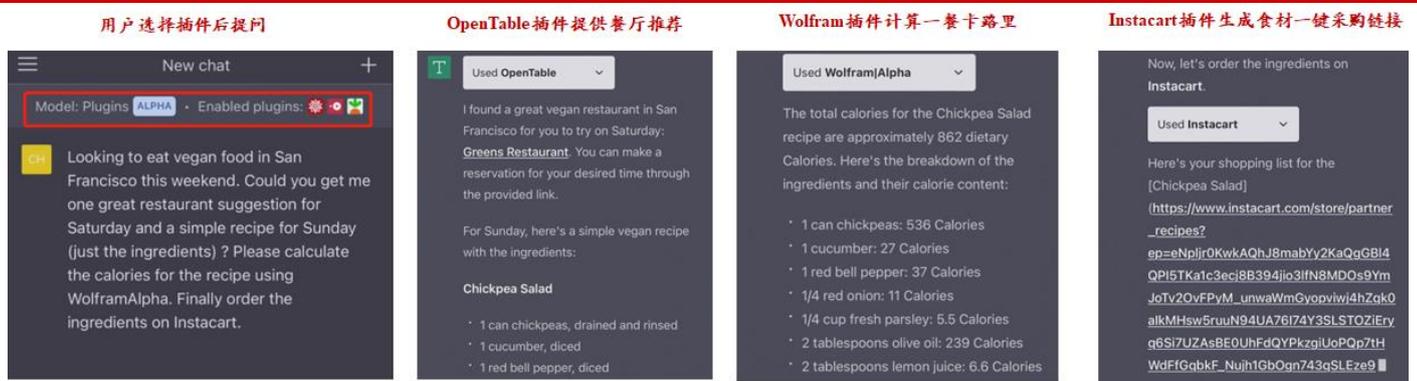
图 15 商店目前已有的插件主要为电商、出行、订餐、资讯等



资料来源：ChatGPT，华西证券研究所

参考官方示例，用户能够使用插件同时调用多个应用的功能，实时解决生活问题。在用户提问“周末在旧金山吃素食，请推荐餐厅、食谱、计算卡路里、采购食材”后，GPT 调用了餐厅推荐、数学计算、生鲜电商插件，在一段对话内完成了这些生活化任务。

图 16 GPT 能够调用多个插件在一段输出内同时解决用户的生活化问题



资料来源：ChatGPT，华西证券研究所

随着外部成熟应用的交互圈层扩大，GPT 与用户日常生活将结合得更加紧密，主流应用将开启“手转 AI”时代。ChatGPT 已经具备了平台雏形，AI 插件商店是 GPT 系统内部的 AppStore，传统应用可能需要重新考虑自身的核心价值和业务模式，将自己的服务无缝融入 ChatGPT 中，用户界面、后端架构以及商业模式均有可能发生改变。

参考移动互联网时代的端转手趋势，主流应用即将展开向 AI 平台的迁移，即迅速适应并布局 GPT 内的 AI 插件，探索 AI 加持下自身应用的新场景，以免被时代淘汰。端转手时代中，各类应用纷纷迁移到移动端，满足用户便捷性需求。如今，GPT 作为 AI 平台，为主流应用提供了更智能、高效的解决方案，将用户体验提升至新高度。

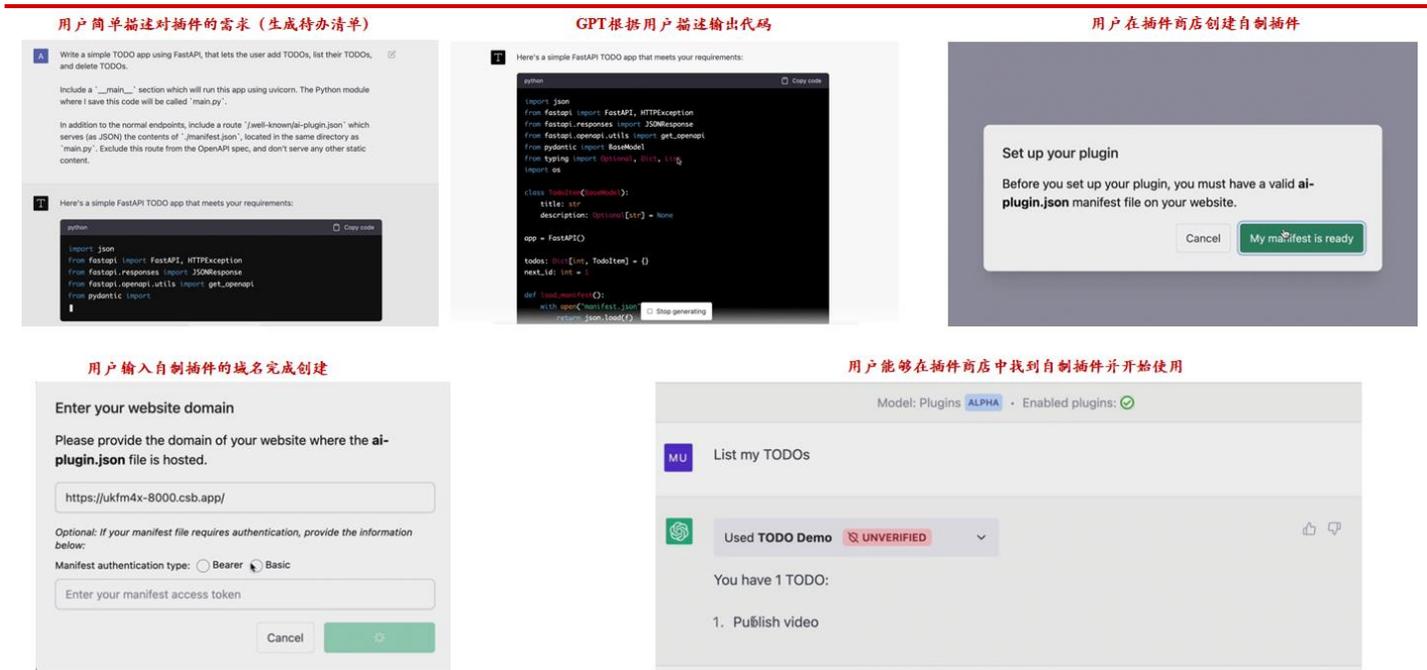
GPT 与外部成熟应用将在生态系统中形成技术与流量的互换关系：成熟应用依赖 GPT 的 AI 技术提高用户体验，而 GPT 则借助这些应用吸引更多用户，网络效应进一步加强了这种相互促进的趋势，使得 GPT 生态价值不断提升。

1.3. 未来的插件生态：开发门槛极大降低，UGC 属性强

PC 和移动互联网时代的开发技术壁垒相对较高。开发者需要具备一定的编程知识和技能，熟悉各种开发工具、编程语言、框架和库。尽管时代发展下开发资源和工具日渐丰富，但仍然开发成熟软件的技术能力依然是一大壁垒。

AI 技术的加持下，用户开发 GPT 插件几乎不需要技术门槛。开发者只需要定义 API、用几句话（不需专业术语）描述对 API 的需求，模型会去完成身份验证、数据处理等工作。我们参考官方演示视频，在创建“待办清单”插件前，用户用正常语言描述对插件的需求（包括三部分：插件的具体用途，插件使用的服务器和模块，插件的清单文件），GPT 能够完成剩下的工作，用户将插件的域名复制到插件商店后，便可以生成并使用插件，整个流程只需三分钟。

图 17 自制一个“代办清单”插件流程只需要三分钟，代码生成、身份验证、交互等工作由 GPT 完成



资料来源: ChatGPT, 华西证券研究所

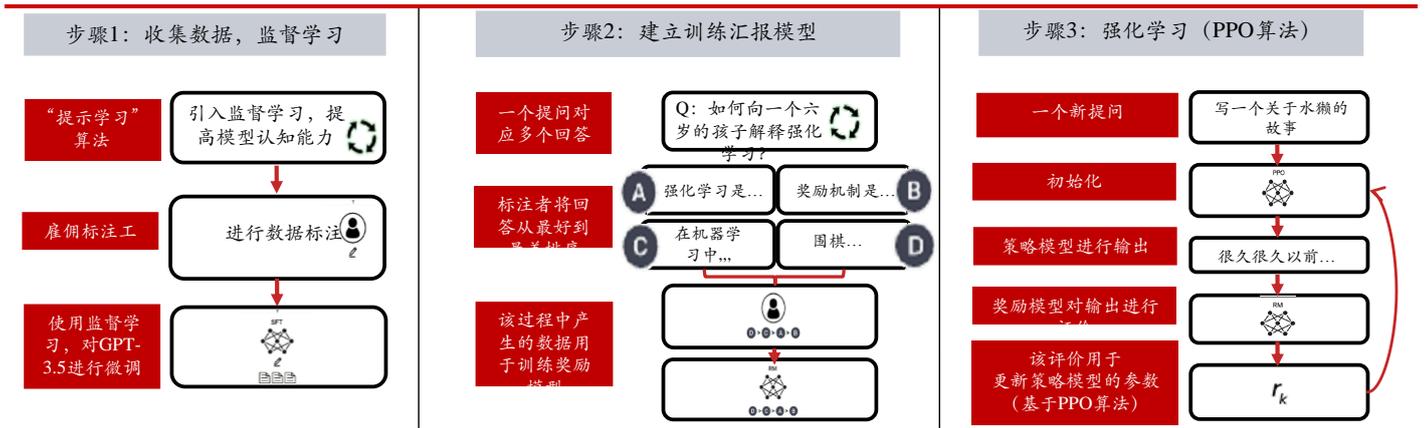
技术门槛的降低意味着 UGC（用户生成内容）回潮，玩法、用户体验和差异化成为关键。在 AI 插件生态中，用户和开发者不必担心复杂的技术实现细节，每个人都可以根据自己的需求创建插件，活跃的社区氛围下用户粘性将提升，插件生态将更加繁荣。对于开发者而言，创意、玩法、差异化体验、与实际场景的融合度等成为了决定自身插件能否取得先发优势的关键，我们认为未来文娱、电商、教育等多个行业格局可能重塑。

2. 从“端转手”到“手转 AI”，国产替代需求更加紧迫

尽管目前海外其他 AI 巨头还未落地 ToC 大模型产品，距离模型的平台化尚有距离，但其旗下的主流应用会否进入微软主导的 GPT 平台仍是未知数。我们维持此前的观点，OpenAI 的成功不涉及底层算法理论的创新，更多是细节处理和工程化的成功。

1) 2022 年 11 月，OpenAI 上线了机器人对话模型 ChatGPT (GPT-3.5)，引入了 RLHF（基于人类反馈的强化学习）：利用人类的标注数据去对 GPT3/GPT3.5 进行有监督训练，针对模型的多个回答进行排序标注，形成奖惩机制，让模型去拟合人的偏好，从而实现了史上最佳的输出效果。ChatGPT 并未实现底层理论的创新，更近于多种前沿算法理论组合，选取了大量的数据，设计了合理的标注流程，并且将这些融合，真正复杂的是这一过程。

图 18 ChatGPT 的技术突破点在于引入了 RLHF（基于人类反馈的强化学习）



资料来源: 公开资料整理, 华西证券研究所

2) 从落地角度来看, ChatGPT 更多实现了工程而非技术上的成功, 即完成了从底层技术到工程落地再到产品的跨越, 在成本、规模和效率之间实现了正确的权衡取舍。从 B 端来看, 此前的 AI 公司需要根据不同的任务训练出不同的模型, 算法很难复用于其他场景。OpenAI 提供了泛化/通用的算法, 让更多的使用者受益并为之消费, 各行业厂商可以利用统一的预训练大模型+微调来直接应对不同的任务, 后续开发工程量大幅下降, 实现了 AI 开发的工业化; 从 C 端来看, ChatGPT 做到了在稳定输出的同时容纳破亿月活用户, 具有低使用门槛、高效、强传播性的特点, 开启了 AI 领域大众化的序章。

若论及学术积累及人才团队, 微软并不具有断层性优势, 当下正是谷歌、Meta 等 AI 框架巨头快速追赶的关键节点, 因此 AI 系统的格局依然风云未定, 微软拥有先发优势, 谷歌、Meta 则具备更多更广阔的应用场景。目前海外 AI 框架领域已经形成 TensorFlow(Google)、PyTorch(Meta) 双寡头格局, 国内主流 AI 框架主要有 PaddlePaddle (百度)、MindSpore (华为)、MegEngine (旷视)、OneFlow 等, 从 Github 指标看, 我国主体推出的 AI 框架中, 华为 MindSpore、百度飞桨引用次数、点赞数、贡献者数量占优。

表 1 Github 社区中主流 AI 框架情况 (2022.1)

地区	排名	框架	代码提交次数	引用数	点赞数	贡献者
海外	1	TensorFlow	124494	86300	163000	3056
	2	PyTorch	43390	14800	53700	2137
	3	Theano	28127	2500	9500	352
	4	CNTK	16116	4400	17100	201
	5	Maneet	11776	6900	19800	868
国内	1	PaddlePaddle	33753	4300	17500	524
	2	MindSpore	37308	514	2700	267
	3	MegEngine	2282	462	4100	32
	4	OneFlow	7621	351	3000	99
	5	Jittor	1266	235	2300	31

资料来源: AI 框架发展白皮书 (2022 年), 华西证券研究所

模型开始成为底层系统, 国产替代变得更加紧迫。相对上一轮“端转手”, 国内厂商在本轮“手转 AI”浪潮中的站位更有优势, 学术积累、人才团队相差不远, 且

根据头部厂商对外披露的信息，除百度“文心一言”外，华为“盘古”大模型、阿里达摩院大模型均在陆续落地中。

➤ **百度：率先落地“文心一言”，搜索场景天然契合**

3月16日“文心一言”发布，从官方demo来看，文心一言具备文学创作、商业文案创作、数理逻辑推算、中文理解、多模态生成能力，存在超预期亮点，但事前录屏降低了其演示的真实性，也并未对外直接开放，多因素导致公众反馈不佳。我们认为尽管上下文理解、语义逻辑、多轮对话方面尚有欠缺，“文心一言”展现了足够的文案创作能力，为B端降本增效的起始逻辑已经明晰，尽管尚未对公众大范围开放，企业用户已经能够申请内测邀请码，若邀请范围持续扩大，飞轮效应将推进“文心一言”表现改善，且优化空间极大。

图 19 “文心一言”生成图片

图 20 “文心一言”生成视频



资料来源：百度，华西证券研究所

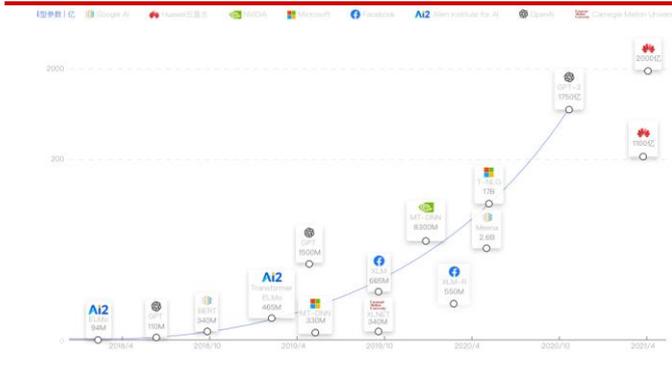
资料来源：百度，华西证券研究所

Bing 的崛起已经证明了人工生成智能作为新介入因素能够对曾经的垄断巨头产生威胁，百度将“文心一言”嵌入搜索具有必要性。此前海内外搜索市场均呈现一家独大的格局，据 statcounter，截至 2022 年 6 月，Google 在全球搜索引擎的市占率达到 92.42%，远超第二名 bing 的 3.45%、第三名 Yahoo! 的 1.32%。但 NewBing 的发布打破了这一稳态，data.ai 数据显示新功能上线当日必应 Bing 应用程序的全球下载量在一夜之间猛增十倍，蹿升至苹果 App Store 应用商店最受欢迎的免费应用榜中的第十位，并使其成为第二大最受欢迎的免费生产力应用，仅次于谷歌邮箱 Gmail。截至 3 月 10 日，Bing 活跃用户已突破 1 亿人，增幅超 600%。我们认为作为国内搜索巨头，百度在搜索领域进行人工生成智能布局已经成为战略上的必需。

➤ **华为：前期积累深厚，近期将发布盘古大模型**

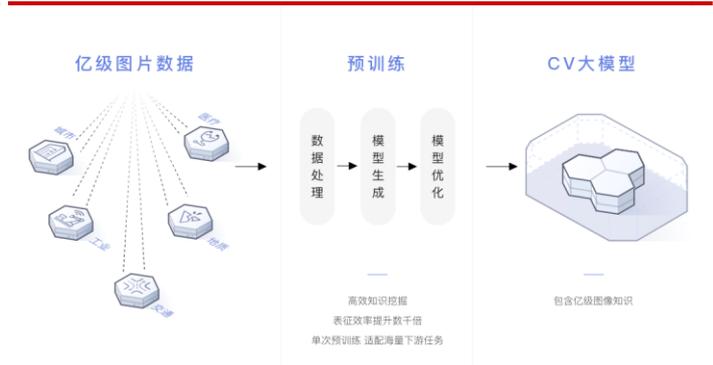
公司从 2020 年开始布局大模型领域，2021 年推出鹏城盘古大模型，成为当时业界首个千亿级、中文 NLP 大模型。即将发布的盘古大模型则涵盖了 NLP 大模型、CV 大模型、多模态大模型、科学技术大模型等多个领域，其中 NLP 大模型对标 ChatGPT，参数量达千亿。同时，华为还推出了业界最大预训练 CV 模型，该模型能够根据模型大小与运行速度需求自适应抽取小模型，助力行业应用快速落地。在大模型产业化方面，华为已经发起了智能遥感开源生态联合体、多模态人工智能产业联合体和智能流体力学产业联合体等。

图 21 盘古 NLP 大模型



资料来源：华为云，华西证券研究所

图 22 盘古 CV 大模型

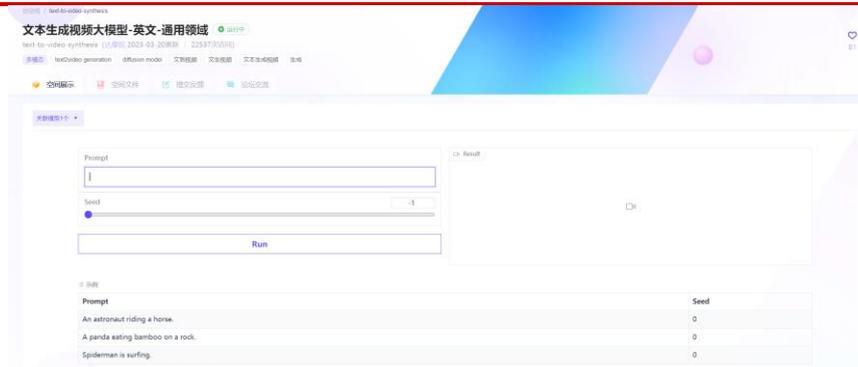


资料来源：华为云，华西证券研究所

➤ 阿里：前期已有大模型技术积累，近期推出生成视频大模型

阿里达摩院在其前期大模型技术积累的基础上，近期推出了多模态大模型，并在 AI 模型社区“魔搭”ModelScope 上线了“文本生成视频大模型”。这一整体模型参数约 17 亿，目前仅支持英文输入。阿里达摩院在 2020 年初启动了中文多模态预训练模型 M6 项目，并于同年 6 月推出了 3 亿参数的基础模型。M6 模型逐步扩展至 2020 年 10 月达到 10 万亿参数规模，成为当时全球最大的 AI 预训练模型。M6 已广泛应用于超过 40 个场景，日调用量上亿，涵盖了电商、制造业、文学艺术和科学研究等多个领域。

图 23 达摩院“文本生成视频大模型”已经上线



资料来源：创空间，华西证券研究所

3.投资建议

模型层：微软主导的 GPT 生态是否适用我国还有待商榷，因此随着大模型平台化、生态化，国产大模型的复现需求变得紧迫。我们在此前报告中已经梳理了模型层复现的难度，判断国内厂商如百度、华为、阿里等复现大模型不会太遥远，国内厂商在本轮“手转 AI”浪潮中的起点相对上一轮“端转手”更有优势。模型层受益标的包括百度、华为（产业链）、阿里、三六零等明确具备大模型落地意愿及能力的厂商，同时追踪腾讯等厂商的相关技术进展。

应用层：GPT 模型的平台化逻辑下，关注两类厂商：

1) 有能力快速接入 GPT 生态，将自身应用插件化的厂商：境外业务占主导、海外用户基础大、有实际应用场景的公司会尽快接入 GPT 生态，先发优势下，我们认为这类厂商的估值将直接受益于 OpenAI 的波浪式创新。

2) 海外映射逻辑下，关注布局游戏、浏览搜索、电商、本地生活等业务，并已有 AIGC 相关技术实践的标的。Meta、谷歌两个 AI 架构巨头还未拿出实际产品，但是其技术积累和微软相比更雄厚，应用场景也更直观，因此通用人工智能领域的想象空间不会止于 SaaS 软件服务，内容消费、社交娱乐等海外巨头更擅长的领域会迅速接棒，而这些领域也是国内厂商的长项。国内头部厂商具备成熟的应用转化及分发能力，产品本身已经建起规模及心智壁垒，且变现路径可以复用，商业模式无需再探索，如要迁移自身应用至 AI 生态内，获客成本远低于新进入者。

4.风险提示

国产大模型落地效果不及预期风险；AI 产品道德及监管风险。

分析师与研究助理简介

赵琳：华西证券传媒行业首席，南开大学本硕。本科毕业后自愿到乡村学校长期支教后担任校长，期间获《中国教育报》头版头条关注报道。2017年硕士毕业后到新时代证券从事传媒行业研究，2019年加盟华西证券。

分析师承诺

作者具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，通过合理判断并得出结论，力求客观、公正，结论不受任何第三方的授意、影响，特此声明。

评级说明

公司评级标准	投资评级	说明
以报告发布日后的6个月内公司股价相对上证指数的涨跌幅为基准。	买入	分析师预测在此期间股价相对强于上证指数达到或超过15%
	增持	分析师预测在此期间股价相对强于上证指数在5%—15%之间
	中性	分析师预测在此期间股价相对上证指数在-5%—5%之间
	减持	分析师预测在此期间股价相对弱于上证指数5%—15%之间
	卖出	分析师预测在此期间股价相对弱于上证指数达到或超过15%
行业评级标准		
以报告发布日后的6个月内行业指数的涨跌幅为基准。	推荐	分析师预测在此期间行业指数相对强于上证指数达到或超过10%
	中性	分析师预测在此期间行业指数相对上证指数在-10%—10%之间
	回避	分析师预测在此期间行业指数相对弱于上证指数达到或超过10%

华西证券研究所：

地址：北京市西城区太平桥大街丰汇园11号丰汇时代大厦南座5层

网址：<http://www.hx168.com.cn/hxzq/hxindex.html>

华西证券免责声明

华西证券股份有限公司（以下简称“本公司”）具备证券投资咨询业务资格。本报告仅供本公司签约客户使用。本公司不会因接收人收到或者经由其他渠道转发收到本报告而直接视其为本公司客户。

本报告基于本公司研究所及其研究人员认为的已经公开的资料或者研究人员的实地调研资料，但本公司对该等信息的准确性、完整性或可靠性不作任何保证。本报告所载资料、意见以及推测仅于本报告发布当日的判断，且这种判断受到研究方法、研究依据等多方面的制约。在不同时期，本公司可发出与本报告所载资料、意见及预测不一致的报告。本公司不保证本报告所含信息始终保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者需自行关注相应更新或修改。

在任何情况下，本报告仅提供给签约客户参考使用，任何信息或所表述的意见绝不构成对任何人的投资建议。市场有风险，投资需谨慎。投资者不应将本报告视为做出投资决策的惟一参考因素，亦不应认为本报告可以取代自己的判断。在任何情况下，本报告均未考虑到个别客户的特殊投资目标、财务状况或需求，不能作为客户进行客户买卖、认购证券或者其他金融工具的保证或邀请。在任何情况下，本公司、本公司员工或者其他关联方均不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告而导致的任何可能损失负有任何责任。投资者因使用本公司研究报告做出的任何投资决策均是独立行为，与本公司、本公司员工及其他关联方无关。

本公司建立起信息隔离墙制度、跨墙制度来规范管理跨部门、跨关联机构之间的信息流动。务请投资者注意，在法律许可的前提下，本公司及其所属关联机构可能会持有报告中提到的公司所发行的证券或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务。在法律许可的前提下，本公司的董事、高级职员或员工可能担任本报告所提到的公司的董事。

所有报告版权均归本公司所有。未经本公司事先书面授权，任何机构或个人不得以任何形式复制、转发或公开传播本报告的全部或部分内容，如需引用、刊发或转载本报告，需注明出处为华西证券研究所，且不得对本报告进行任何有悖原意的引用、删节和修改。