

从阿里/商汤/华为大模型看应用趋势

华泰研究

2023年4月20日 | 中国内地

专题研究

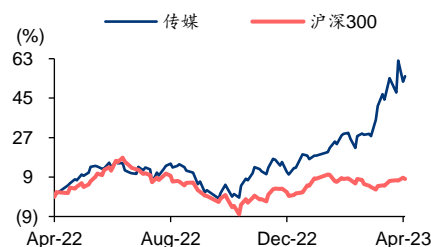
传媒

增持 (维持)

研究员 朱璐
 SAC No. S0570520040004 zhujun016731@htsc.com
 SFC No. BPX711 +(86) 10 6321 1166

联系人 王星云
 SAC No. S0570121100014 wangxingyun@htsc.com
 +(86) 755 8249 2388

行业走势图



资料来源: Wind, 华泰研究

国产大模型频发, 数据集/IP、视频/虚拟人、游戏等领域迎机遇

近期国内大模型频繁发布, 如阿里发布大语言模型“通义千问”, 商汤推出“商量”, 华为的盘古大模型则聚焦 AI for Industry。我们认为, 拥有稀缺数据集和版权 IP 的公司有望首先受益于大模型的普及应用, 后续影响同样将延伸到视频/虚拟人、游戏制作等领域。数据集方面, 高壁垒的数据集或将受益于大模型的训练; IP 版权方面, 作为数据训练源头的版权 IP 将迎来海量需求, 行业产能升级也将反哺版权 IP 创作; 视频/虚拟人方面, AI 有望提升制作效率降低成本, 同时虚拟人的互动性有望增强; 游戏方面, 大模型可在 NPC 聊天内容、剧情大纲、场景生成等领域助力游戏生产。

阿里发布通义大模型, 迎接智能化时代

阿里大语言模型“通义千问”, 具备多轮对话、文案创作、逻辑推理、多模态理解、多语言支持等功能, 以及外部增强 API。未来阿里所有产品未来将接入大模型全面升级, 目前钉钉、天猫精灵等产品已率先接入通义千问测试, 高德地图、饿了么、盒马、优酷、淘票票等产品也将有序接入通义千问大模型。同时阿里宣布开发企业专属大模型, 支持 Web 界面与专属 API, 满足企业个性化需求。我们认为, 随着阿里大模型的发布, “让算力更普惠”、“让 AI 更普及”目标的提出, 研发门槛将大大降低, 将加速智能化时代的到来。

商汤发布日日新大模型体系, 推动 AGI 的到来

商汤“日日新”大模型体系由商汤 AI 大装置 SenseCore 提供算力支持, 包括自然语言大模型“商量”、中文医疗语言大模型“商量·大医”、编程助手“商量·AI 代码助手”三个平台, 并进一步推出了四款生成式 AI 应用, 包括“秒画 SenseMirage”文生图创作平台、“如影 SenseAvatar”AI 数字人视频生成平台、“琼宇 SenseSpace”3D 场景生成平台和“格物 SenseThings”3D 内容生成平台。我们认为, 通过不断完善“日日新”大模型体系, 并进一步提供 MaaS 服务、开放 API 等商业落地服务, 下游应用有望迎来较快发展。

华为盘古大模型聚焦 AI for Industry, 赋能千行百业应用落地

华为盘古大模型分为基础模型 L0 (NLP、CV、科学计算三大模型), 行业模型 L1 (矿山、气象、电力、药物分子、金融等) 以及细分场景模型 L2 三个层级。我们观察到, 盘古大模型聚焦 AI for Industry, 过去几年的 AI 项目应用已超过 1,000 个, 其中 30% 用户用于客户的核心生产体系中, 平均推动客户盈利提升 18%, 且目前在工业 B 端领域发力效果尤为显著。盘古大模型具备低门槛 AI 开发、泛化性能高、高效样本标注、精准度高、解决碎片化问题五大核心优势, 能够有效解决传统 AI 开发难题。我们认为, 华为盘古大模型有望推动人工智能开发从“作坊化”到“工业化”升级。

风险提示: AIGC 技术发展不及预期, 政策监管风险, 商业化存在不确定性; 本研报中涉及到未上市公司或未覆盖个股内容, 均系对其客观公开信息的整理, 并不代表本研究团队对该公司、该股票的推荐或覆盖。

正文目录

阿里巴巴通义大模型	3
全栈布局 AI 技术体系，MaaS 成为重要层级.....	3
推出通义千问语言模型，后续将陆续接入阿里内部产品.....	4
开发企业专属大模型，与 OPPO 等多家企业开展合作.....	8
商汤日日新大模型	9
SenseCore 大装置：为大模型提供充足的算力支持.....	9
商量大语言模型赋能专业知识、医疗、编程等领域.....	9
配套生成式 AI 应用，驱动垂直场景降本增效.....	11
大模型商业变现路径：MaaS 模式+开放 API 接口.....	13
华为盘古大模型	15
三层级布局大模型.....	15
大模型聚焦 AI for Industry，驱动行业智能升级.....	15
大模型推动 AI 开发从“作坊化”到“工业化”升级.....	17
从大模型发展看应用层机会	18
数据集/IP：优质数据或成宝贵训练资源，有望反哺 IP 生产.....	18
视频/虚拟人：缩短制作周期，提升交互能力.....	19
游戏：赋能游戏素材生成，促进玩法创新.....	20
报告提及公司.....	21
风险提示.....	21

阿里巴巴通义大模型

全栈布局 AI 技术体系，MaaS 成为重要层级

阿里全栈布局 AI 技术体系，形成 IaaS、PaaS 和 MaaS 三层架构。阿里（阿里巴巴）表示，在 IaaS（Infrastructure as a Service）层，公司为 AI 设计了云基础设施，包括计算的模块、高效的网络及储存，如灵骏智能计算集群与弹性计算 ECS 集群。公司的 PaaS（Platform as a Service）层提供了丰富的大数据及机器学习产品，能够从数据清洗开始帮助开发者训练模型。MaaS（Model as a Service）层包括基础大模型/通义大模型、企业专属大模型、魔搭社区、API 服务等。

图表1：阿里云实现 IaaS、PaaS 和 MaaS 三层架构



资料来源：阿里云官网、华泰研究

阿里云追求让算力更普惠、让 AI 更普及。1) 让算力更普惠：“普”就是要推动低代码及 Serverless（无服务器）发展，提升开发者服务，降低开发门槛；“惠”就是加大软硬一体技术开发，提高规模化运营效率，持续释放技术红利。阿里表示，过去十年单位算力成本下降 80%，存储成本下降 88%，公司希望未来的算力成本变成今天的十分之一乃至百分之一。2) 让 AI 更普及：公司希望在云基础设施平台上，有阿里的通义大模型，也有阿里云跟高校、科研院所、各行各业合作的基础大模型，平台不仅支撑基础大模型的训练和服务，也为各行业的专属大模型提供保障。

图表2：阿里云追求让算力更普惠



资料来源：阿里云官网、华泰研究

MaaS 已成重要层级，魔搭社区模型总数达 800+。2022 年 11 月，阿里巴巴在云栖大会上首次推出 MaaS 的概念，即以云为基础，以模型为中心。MaaS 的核心是将模型作为生产的重要元素，有效支撑模型的生命周期，模型研发、数据清洗、模型训练、模型测试等能够进入统一模型网站，开发者能快速查找并使用模型，降低模型使用门槛。同时使用模型做二次开发，解决各行业问题。魔搭（ModelScope）作为 AI 开源社区，提供视觉、语音、自然语言等各个领域的 SOTA 预训练模型，自阿里巴巴在 2022 年 11 月推出以来，魔搭目前已拥有上百万开发者，模型累计下载次数超 1600 万，模型总数达 800+ 个。

图表3：MaaS 开放平台魔搭社区



资料来源：阿里云官网、华泰研究

推出通义千问语言模型，后续将陆续接入阿里内部产品

阿里云智能首席技术官周靖人在 2023 阿里云峰会上正式宣布推出大语言模型通义千问，支持多轮交互及复杂指令理解、多模态融合，以及外部增强 API。

通义千问具备多轮对话、文案创作、逻辑推理、多模态理解、多语言支持等功能。通义千问功能包括：**1) 多轮对话**：能够了解用户表达的意图，通过上下文语义理解回答问题；**2) 文案创作能力**：能够续写小说、编写邮件、实现企业策划，可以极大增加工作效率；**3) 逻辑推理**：能够完成简单数学题及相关代码的编写，为开发者在开发流程上提效；**4) 多模态理解**：能够将视觉、视频的知识融合在语言体系中；**5) 多语言支持**：实现多语言的交互及转接。

图表4：通义千问具备多个功能



资料来源：阿里云官网、华泰研究

“通义千问”可在多个场景投入应用。1) 智能办公方面，通义模型可调用差旅接口推荐差旅产品、调用导航接口规划交通路线、在视频会议中自动生成会议摘要、撰写邀请函并制作海报等；2) 智能购物方面，通义模型能够根据用户需要回答问题并推荐具体产品、提供活动策划方案及场地推荐；3) 智能居家方面，通义千问能够根据用户需要，实现个性化故事生成、个性化歌单、个性化菜谱等。

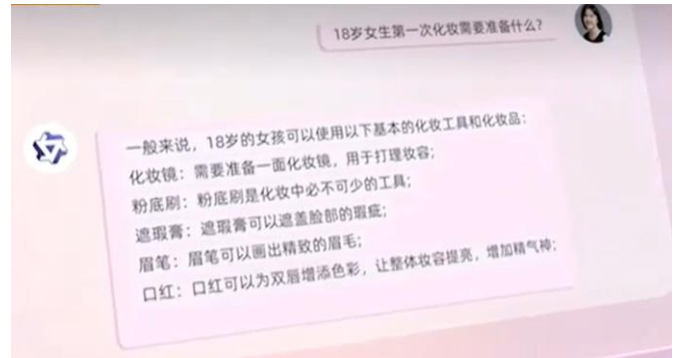
通义模型支持多种接入调用方式。基于灵骏平台，通义模型支持各种模型一键部署云上服务，开发者可自动获得 API Key 进行模型微调 and 指令化。

图表5：通义千问调用差旅接口为用户推荐机票及酒店



资料来源：阿里云官网、华泰研究

图表6：通义千问为用户解答问题并推荐产品



资料来源：阿里云官网、华泰研究

阿里云 CEO 张勇表示，在 AI 时代，所有产品都值得用大模型重新升级。阿里云计划未来先从内部生态开始，将所有产品接入“通义千问”大模型，进行全面升级，让各个业务去更好的独立面对市场，经受挑战，提升其智能化水平，然后再面向社会企业开放，帮助更多企业用上大模型，让每家企业都能基于“通义千问”打造属于自己的大模型，阿里云期望与合作企业共同探索数字化未来。

图表7：阿里所有内部产品都将接入通义千问大模型



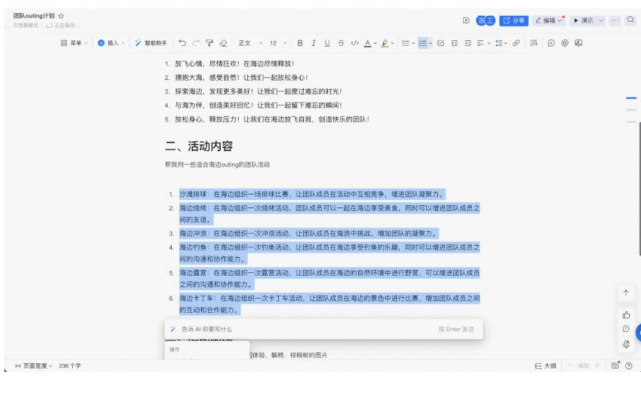
资料来源：阿里云官网、华泰研究

目前钉钉、天猫精灵等产品已率先接入通义千问测试，高德地图、饿了么、盒马、优酷、淘票票等产品也将有序接入通义千问大模型。

钉钉接入通义千问模型后，可以在钉钉文档、钉钉会议、钉钉群聊、钉钉小程序中实现近 10 项自动化功能，包括：

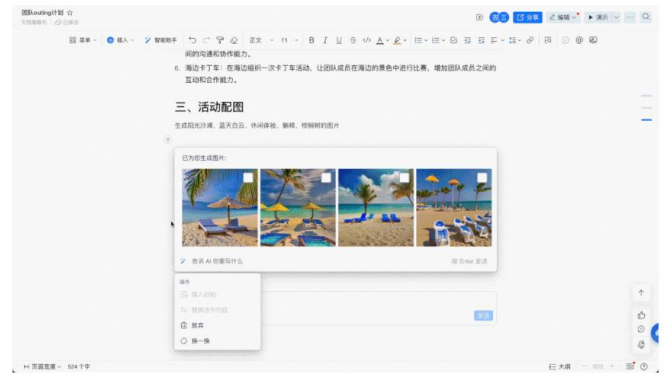
1) 钉钉文档：接入通义千问模型后，用户在文档中输入指令，即可让钉钉完成写作、撰写邮件、生成策划方案、生成图片等一系列工作。

图表8：钉钉文档在接入通义千问模型后可生成营销策划方案



资料来源：阿里云官网、华泰研究

图表9：钉钉文档接入通义千问模型后可根据指令生成图片



资料来源：阿里云官网、华泰研究

2) 钉钉群聊：接入通义千问模型后，钉钉可为用户自动生成未读聊天的内容摘要，帮助用户快速了解上下文信息，解决了手动翻阅效率低下问题。同时，钉钉还可以为用户在群聊中达成的共识自动生成待办事项。

图表10：钉钉可为用户总结群聊的内容摘要



资料来源：阿里云官网、华泰研究

图表11：钉钉可为用户在群聊中达成的共识生成待办事项



资料来源：阿里云官网、华泰研究

3) 钉钉会议：接入通义千问模型后，钉钉可以在会议过程中实时生成字幕，用户可以看到发言人以及其所讲的内容。会议后，钉钉还能够自动生成会议重点摘要与待办事项。

图表12：钉钉可在会议中实时生成字幕并能够总结会议纪要



资料来源：阿里云官网、华泰研究

4) 拍照生成小程序：接入通义千问模型后，用户无需写代码，只需在钉钉中上传一张功能草图，钉钉即可立刻生成一款应用小程序。

图表13：钉钉可自动生成订餐小程序



资料来源：阿里云官网、华泰研究

接入通义千问模型后，天猫精灵变得比以前更聪明，在知识、情感、个性、记忆能力方面都有显著的提升，智能化水平大幅跃升。它可以根据用户需求和交流场景随时生成内容、切换话题，可以像真人一样自由对话。例如，用户可以：1) 要求天猫精灵与自己一同创作“宇宙大爆炸”的新故事；2) 让天猫精灵回答刁钻问题；3) 要求天猫精灵合成适合跑步的歌单；4) 与天猫精灵谈论人生。

图表14：阿里巴巴天猫精灵



资料来源：阿里云官网、华泰研究

开发企业专属大模型，与 OPPO 等多家企业开展合作

开发企业专属大模型，支持 Web 界面与专属 API。阿里巴巴表示，阿里云希望能够帮助更多企业用上大模型，让每家企业都能基于“通义千问”，具备自己行业能力的大模型。首先，建立企业专属数据空间，企业能够将数据加入数据空间中，可接入各种形式的文档及云上数据库。之后，模型能够自动生成企业专属大模型。最后，模型能够自动学习企业数据空间中的信息，形成符合企业使用习惯、行业属性的大模型。所有模型生成都能够自动化地生成，生成的模型将自动放置在阿里云上进行推理配置。企业也可以嵌入 Web 系统开发，也可以将 API 接入自己的业务系统。

目前通义千问模型已与多家企业展开合作探索。首批接入的合作企业包括 OPPO、安迪斯智能云、太平洋保险、吉利汽车、上汽致以、奇瑞新智能、豪墨自行、太古可口可乐、波司登、顾家家居、V1、酷家乐等。阿里表示，未来将基于通义千问将持续创新，做好语言服务。

图表15：阿里云开发企业专属大模型



资料来源：阿里云官网、华泰研究

商汤日日新大模型

SenseCore 大装置：为大模型提供充足的算力支持

商汤“日日新 SenseNova”大模型体系的基础算力底座是商汤 AI 大装置 SenseCore，其背靠商汤人工智能计算中心（AIDC），目前有 27000 块的 GPU 在运行，能够输出 5000p 的总算力，500p 的国产化算力，是亚洲目前为止最大的智算中心之一，可以同步支持 20 个千亿参数超大模型训练，为商汤日日新大模型体系提供了充足的算力支持。基于大装置的能力，商汤目前已构建了计算机视觉、自然语言处理、AI 内容生成、多模态、决策智能等多个领域的大模型。

图表16：商汤 AI 大装置 SenseCore 为大模型体系提供算力支持



资料来源：商汤技术交流日、华泰研究

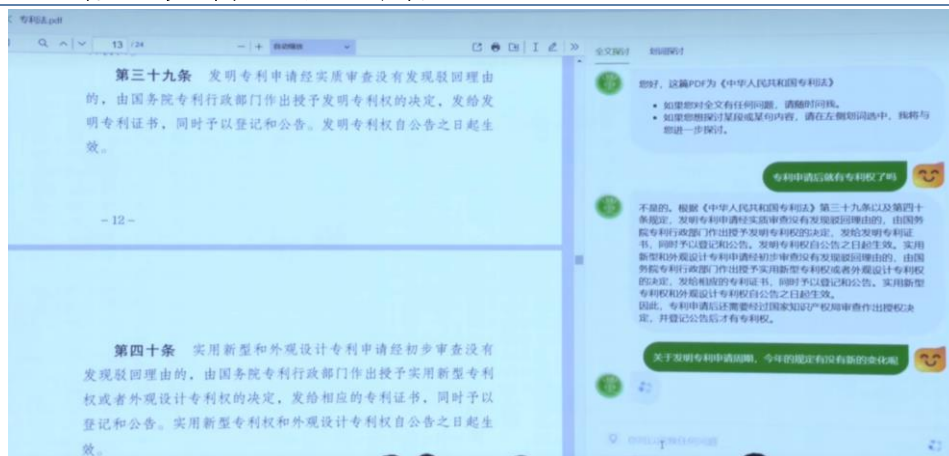
商量大语言模型赋能专业知识、医疗、编程等领域

商汤目前已基于 AI 大装置，推出了自然语言大模型“商量”、中文医疗语言大模型“商量·大医”、编程助手“商量·AI 代码助手”三个应用平台，赋能专业知识、医疗、编程等垂类场景。

1) 自然语言大模型“商量”（SenseChat）

自然语言大模型“商量”拥有 1800 亿参数，具备超长文本知识理解、文本生成、内容创作、多轮问答、逻辑推理等能力。它通过与用户进行多轮互动来解决复杂问题、提供定制化建议、辅助创作文本，且自带知识自动及时更新模块，从而生成更可信、更准确和更安全的信息。例如，用户可以通过与“商量”进行对话，让它辅助创作一段广告、一份邀请函，“商量”也能够接入财务、法律等专业知识领域，对用户上传的专业领域长文本 PDF 进行理解和对话。

图表17：“商量”能够理解专业领域的长文本内容

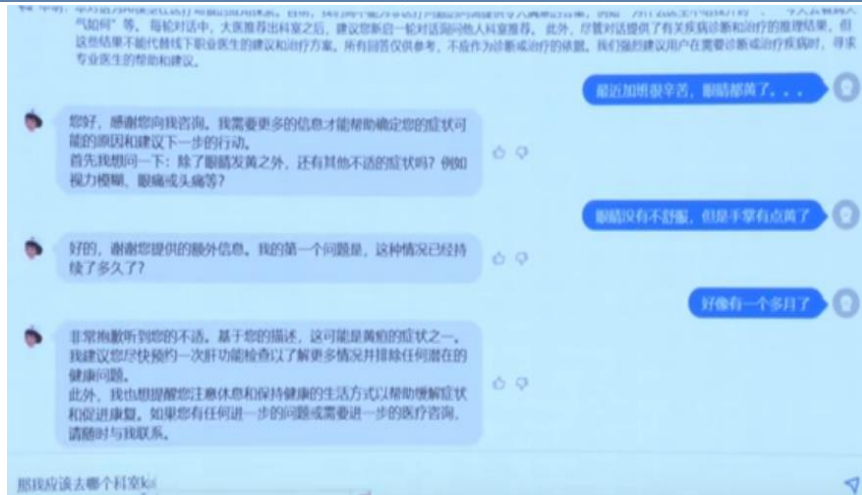


资料来源：商汤技术交流日、华泰研究

2) 中文医疗语言大模型“商量·大医”

中文医疗语言大模型“商量·大医”基于海量医学知识和真实医患互动对话数据打造，它能够通过多轮对话辅助支持导诊、问诊、健康咨询、辅助决策等多场景，从而持续赋能医疗领域，提升医院诊疗效率，为患者打造更好的服务体验。据商谈科技展示的案例，“商量·大医”已落地新华医院，能够担任健康咨询助手，为用户解答熬夜会带来身体变化的原因，并通过引导给用户就医挂号建议。

图表18：“商量·大医”能够提供在线问诊服务



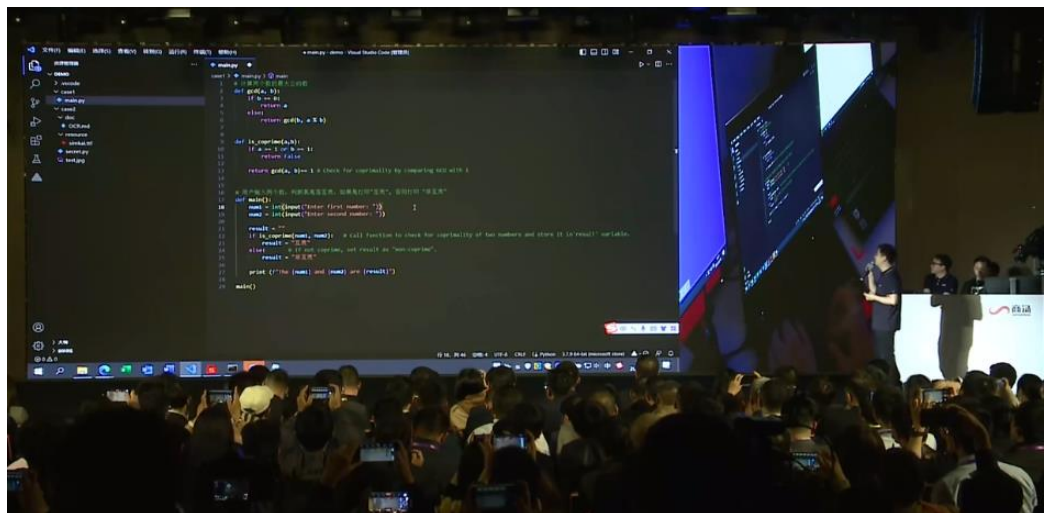
资料来源：商汤技术交流日、华泰研究

3) 编程助手“商量·AI代码助手”

编程助手“商量·AI代码助手”具备代码补全、代码扩写、代码翻译、代码重构、代码修正、注释生成代码、复杂度分析、测试用例生成等功能，并能支持多语言的中英文编程。“商量·AI代码助手”创造了软件开发新范式，改变了代码编写领域原有的二八定律，即80%由人工生成，剩下的20%由AI完成，构建了新的二八定律：代码=80%AI生成+20%人工，这将有助于用户更高效的编写和调试代码，提升工作效率，简化开发流程。据商汤内部实测，使用“商量·AI助手”后，员工编写代码的效率提升62%，HumanEval测试集的一次性通过率达到了40.2%，数据表现优于Copilot。

据商汤科技展示的案例，AI自主编写了代码，通过递归辗转相除法计算出了两个数的最大公约数，并进一步判断了两个数是否互为质数。

图表19：“商量·AI代码助手”能够编写代码计算数学问题



资料来源：商汤技术交流日、华泰研究

配套生成式 AI 应用，驱动垂直场景降本增效

基于 SenseCore 大装置及“商量”大语言模型，商汤科技还推出了“日日新”大模型体系下的生成式 AI 应用，包括“秒画 SenseMirage”文生图创作平台、“如影 SenseAvatar”AI 数字人视频生成平台、“琼宇 SenseSpace”3D 场景生成平台和“格物 SenseThings”3D 内容生成平台。

#1 “秒画 SenseMirage”文生图创作平台

“秒画 SenseMirage”文生图创作平台融合了超 10 亿参数的商汤自研文生图模型，能够根据用户提示词生成图片，其具备以下特点：1) 推理速度快。秒画由单卡 A100 支持，能够在 2 秒内生成 1 张 512K 分辨率的图片；2) 可自训练。LoRA 小模型搭载在大模型上，用户可通过自身需求，通过图片实例训练生成大模型没见过的东西，5 分钟内就能定制属于自己的 LoRA 小模型，使结果更精确；3) 提供生成模型开源社区。秒画汇聚了 1 万+开源模型，能够为创作者提供加速内容生产创作工具的生态平台；4) 提供 B 端 API 服务。

根据商汤科技展示的案例，原有的文生图模型难以理解“港风美女”这一提示，而上传 20 张左右的香港明星照片后，对模型进行增量训练，短时间内秒画就可以生成图片。

图表20：“秒画 SenseMirage”文生图创作平台能够根据提示生成图片



资料来源：商汤技术交流日、华泰研究

#2 “如影 SenseAvatar”AI 数字人视频生成平台

“如影 SenseAvatar”AI 数字人视频生成平台具备全栈智能创作能力，能够实现 AI 数字人动作表情生成、AI 文案生成、AI 跨语言文稿生成、AI 素材生成等功能，实现在直播、营销视频等场景的落地。用户仅需提供一段 5 分钟的真人视频素材，如影就可以生成出来声音及动作自然、口型准确、多语种精通的数字人分身。

如影具备以下特点：1) 风格种类丰富。如影可生成 2D、3D、卡通等数字人形象；2) 集成了文本生成、文生图、实时问答等多种工具。如影可以根据用户的简单提示生成文案，然后进一步根据文案自动生成数字人视频，并可以依据用户需求提供不同的语言版本。

据商汤科技示例，用户先在大模型中简单描述与故宫联合开发的文创产品，如影就可以生成一段视频，并配上视频文案。

图表21：“如影 SenseAvatar” AI 数字人视频生成平台界面



资料来源：商汤技术交流日、华泰研究

#3 “琼宇 SenseSpace” 3D 场景生成平台

“琼宇 SenseSpace” 3D 场景生成平台具备城市级大尺度的空间重建生成能力，它能够：

- 1) 快速生成城市级 3D 场景。琼宇可以在 2 天内生成 100 平方公里的场景，缩减原有 10000 人·天的工作量；
- 2) 高还原场景细节。琼宇可以厘米级重建精度，还原场景的真实细节和光照效果，支持实时渲染和互动，例如它可以渲染出墙面的反光、地板的灯光。琼宇 3D 场景生成平台可应用于影视创作、建筑设计、商品营销、数字孪生管理运营等垂直领域。

图表22：“琼宇 SenseSpace” 3D 空间生成平台



资料来源：商汤技术交流日、华泰研究

#4 “格物 SenseThings” 3D 内容生成平台

“格物 SenseThings” 3D 内容生成平台具备物体的精细化 3D 生成能力，可以还原室内场景的物体复杂的几何结构、纹理、材质、光泽等信息，做到实时高逼真渲染。传统 3D 建模具备难点：1) 复杂物体建模会牵连到背景；2) 有光泽的物体无法分辨材质。但格物系统能够：1) 复刻复杂结构物体；2) 精准复刻光照；3) 完美还原材质。通过扫描快速建立 3D 物体模型，格物平台综合扫描效率提升了 400%，成本降低了 95%。

格物能够应用于空间的创意设计中，可以做家装、影视作品的嵌入、综艺视频中的物体摆件等，也可以运用到数字人直播间的 3D 场景选择步骤中。

图表23：“格物 SenseThings” 3D 内容生成平台


资料来源：商汤技术交流日、华泰研究

大模型商业变现路径：MaaS 模式+开放 API 接口

基于商汤 AI 大装置 SenseCore，商汤面向客户提供了多种 MaaS（Model as a Service）服务，包括：

- 1) 自动化数据标注。**相较于人工数据标注，基于预训练大模型的自动化数据标注可实现近百倍的效率提升；
- 2) 自定义大模型训练及模型增量训练。**大模型并行训练和模型增量训练服务能够帮助客户快速利用自有数据训练模型，在预训练大模型上对某垂域行业模型进行开发，生产出各行各业的自定义模型；
- 3) 模型推理部署。**模型推理部署服务可降低用模型提供服务的成本，并将大模型推理效率提高 100% 以上；
- 4) 开放模型及 AI 工具链。**全面赋能客户提升开发效率。

目前商汤大装置已有超 7000 张 GPU 算力对外服务 8 家大型客户，用来训练客户超过千亿规模的大参数模型，客户包括科研机构、头部商业银行、头部游戏公司、头部互联网公司、明星创业公司等。

图表24：商汤面向客户提供多种 MaaS 服务


资料来源：商汤技术交流日、华泰研究

同时，基于商汤“日日新”大模型体系，商汤科技开放了 API 接口，方便用户调用各项 AI 技术以实现二次开发，包括：

1) 自然语言生成的 API。自然语言生成服务支持中文的多轮对话，具备超长文本的理解能力，可以不断学习进化。

2) 图片生成的 API。图片生成 API 可以支持文生图、图生图，能够支持 6K 的高清图像，能够生成不同风格图像，用户可根据需要调用 API 进行自主训练。

3) 视觉的通用感知任务和标注的 API。视觉标注服务由商汤科技研发的数据标注平台“明眸”提供服务，明眸内置 12 个通用大模型和行业专用大模型，支持智能驾驶、智慧交通、智慧城市等多种场景的 2D 分类、检测和 3D 检测的智能标注；相比传统人工标注和小模型标注模式，明眸具有标注效果好，效率高，成本低等核心优势，能够对一些密集的车辆、行人等颗粒度较细的图像准确检测。通过开放标准 API，客户能够从底层调度商汤大装置海量算力，实现大规模标注。

图表25：明眸提供自动数据标注服务



资料来源：商汤技术交流日、华泰研究

华为盘古大模型

三层级布局大模型

华为盘古大模型分为三个层次，包括基础模型 L0，行业模型 L1 以及细分场景模型 L2。

1) 基础模型 L0。基础模型基于一站式 AI 开发平台 ModelArts 进行加速优化训练，包括 NLP 大模型、CV 大模、科学计算大模型。其中 NLP 大模型具备文本生成、内容理解能力，CV 大模型具备图像分类、分割、检测能力，盘古科学计算大模型涵盖气象预报、药物分子优化、海浪预测等能力。

2) 行业模型 L1。行业模型是在基础模型上导入行业数据进行训练后衍生而成的模型，包括矿山、气象、药物分子、电力、海浪、金融等行业大模型。

3) 细分场景模型 L2。细分场景模型是在行业模型的基础上结合行业细分场景，对数据进行微调和部署得到的推理模型，包括金融 OCR、金融违约风险识别、电力质检、海浪预测、小分子优化、空调空气质量优化等场景模型。

图表26： 三次布局盘古大模型



资料来源：华为官网、华泰研究

大模型聚焦 AI for Industry，驱动行业智能升级

盘古大模型聚焦 AI for Industry，致力于做好行业应用，赋能煤矿、水泥、电力、金融、农业、国家云等 B 端行业。其中盘古 NLP 大模型可应用于智能文档搜索、智能 ERP、小语种大模型领域，具体场景覆盖类案检索、企业财务异常检测、阿拉伯语大模型；盘古 CV 大模型应用于工业质检、物流仓库监控、时尚辅助设计，具体场景覆盖铁路 TFDS、物的银行、门店半定制设计等；盘古科学计算大模型适用于气象预报、海浪预测等领域，现已推出盘古气象大模型及盘古药物分子大模型。过去几年，华为云的 AI 项目应用已超过 1,000 个，其中 30% 用户用于客户的核心生产体系中，平均推动客户盈利提升 18%。

盘古 CV 大模型具备 30 亿参数及 10 亿级图像，是业界最大的 CV 预训练模型，兼顾判别与生成能力，小样本学习性能领先，在 ImageNet 线性分类领域上为业界第一。据华为展示的案例，盘古 CV 大模型在矿山及铁路领域赋能效果显著：

1) **华为云盘古矿山大模型**。盘古矿山大模型基于海量矿山数据自主学习而成，可在煤矿综采场景、煤矿主运场景、煤矿作业场景得到运用，帮助降低劳动强度，减少安全风险。在综采场景中，矿山大模型可通过 5F+AI 全景视频拼接出 40 米综采面画卷，保证安全生产；在主运场景中，矿山大模型对异物识别精度达 98%；在作业场景中，矿山大模型可用于作业序列智能检测系统，对动作规范识别准确率达 95%，可减少 90% 以上的安全事故。

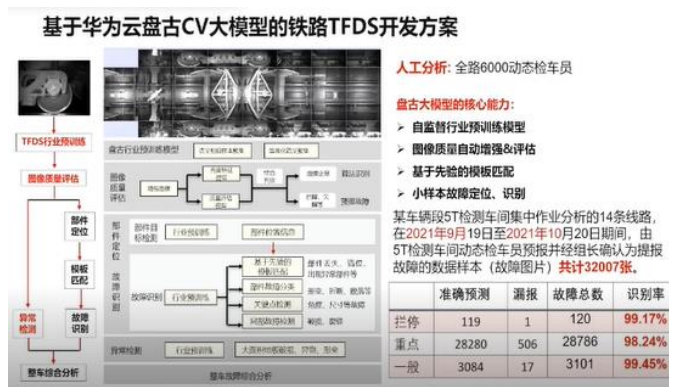
2) **基于华为云盘古 CV 大模型的铁路 TFDS (货车故障轨边图像检测系统) 开发方案**。与传统的人工分析方案相比，大模型降低了检测成本，减轻了动态检车员的工作量，对于故障样本检测的识别率可超 98%。

图表27: 盘古矿山大模型



资料来源：华为官网、华泰研究

图表28: 基于盘古 CV 大模型的铁路 TFDS 开发方案



资料来源：华为官网、华泰研究

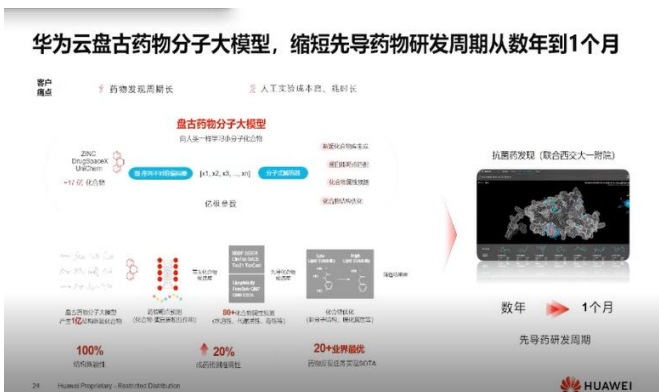
同时，盘古科学计算大模型也在气象、药物分子以及空气质量优化三个领域得到具体应用：

1) **盘古气象大模型**。盘古气象大模型是全球首个精度超过传统预报方式的 AI 模型，能够在秒级时间内给出未来 7 天的天气，且在 1 小时-7 天的预测精度均超过欧洲气象局，预测速度相比欧洲气象局提升 10,000 倍，台风轨迹预测精度提升 20%。

2) **盘古药物分子大模型**。盘古药物分子大模型可以将先导药物研发周期从数年缩短至 1 个月，解决了药物发现周期长、人工实验成本高、耗时长等行业痛点。

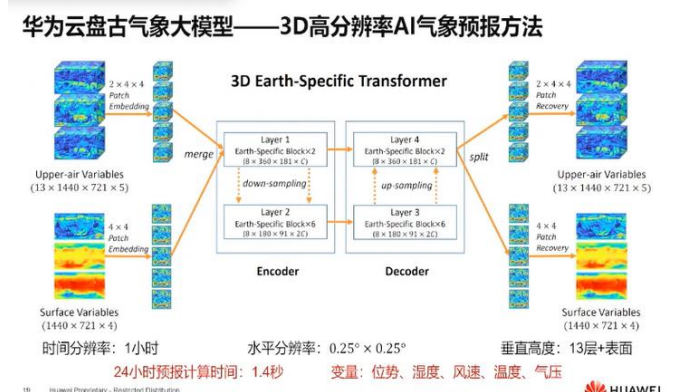
3) **空气质量优化**。盘古大模型赋能空气质量优化，能让电力消耗降低 15% 以上，对甲醛预测精度大于 85%，对二氧化碳预测精度大于 95%。

图表29: 盘古药物分子大模型



资料来源：华为官网、华泰研究

图表30: 盘古气象大模型



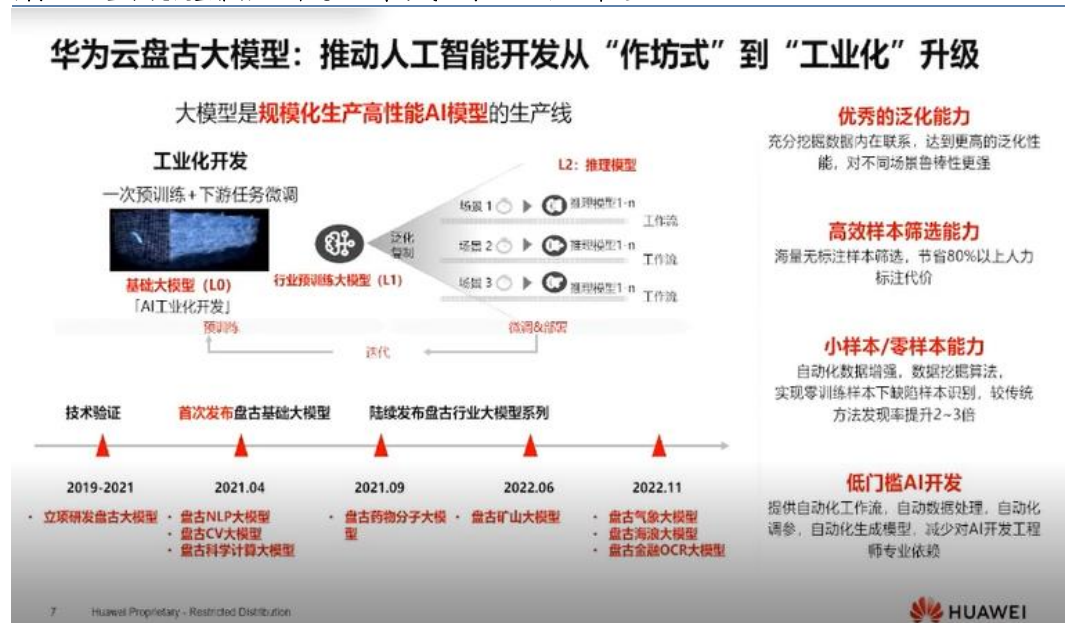
资料来源：华为官网、华泰研究

大模型推动 AI 开发从“作坊化”到“工业化”升级

随着人工智能逐渐渗透到千行百业的核心生产系统，未来人工智能发展趋势包括：**1) 从小模型到大模型**。传统 AI 时代下，一个场景对应一个模型，存在模型参数量小，泛化性差，模型维护困难、行业人短缺等问题，但过去 10 年内，AI 算法的算力需求提升了 40 万倍，模型复杂度持续爆发，而大模型凭借着资本优势及技术优势，可能收编高度定制化的小模型，导致市场向大公司集中；**2) AI+传统科学计算**。工业、气象、能源、生物医学等领域都受到 AI 的深刻影响，AI 能够为传统科学计算带来新思路、新工具和新方法，而传统科学计算也能为 AI 带来更严密的科学指导。

而盘古大模型具备低门槛 AI 开发、泛化性能高、高效样本标注、精准度高、解决碎片化问题五大核心优势，不仅能够解决传统 AI 时代下小模型存在的难题，也能够规模化高性能 AI 模型的生产线，推动人工智能开发从“作坊化”到“工业化”升级，顺应了 AI 发展趋势，将成为未来 AI 生态的核心。

图表31：盘古大模型推动 AI 开发从“作坊式”到“工业化”升级



资料来源：华为官网、华泰研究

从大模型发展看应用层机会

数据集/IP：优质数据或成宝贵训练资源，有望反哺 IP 生产

作为高壁垒资源，数据集有望受益于大模型的普及应用。OpenAI 的主要训练数据包括维基百科、书籍、期刊、Reddit 链接、Common Crawl 和其他数据集等，大量丰富的数据集资源是 OpenAI 打造模型的基础。2023 年 4 月 11 日，网信办公布《生成式人工智能服务管理办法（征求意见稿）》明确指出，鼓励优先采用安全可信的数据资源；训练数据应合法，不侵犯知识产权，应保证其真实性、准确性、客观性和多样性。在各大厂纷纷推出自有大模型的情况下，优质的数据集愈发成为宝贵资源。据深交所互动易，国内多家模型公司与中文在线正在就采购公司中文数据进行合作磋商；作为在图书出版领域掌握大量中文独家语料数据的出版机构，中信出版已就运用 AIGC 积累数字资产、创新收入模式方面认真研究论证，组织专项技术和业务团队，积极孵化创新项目。

图表32：中信出版掌握的中文独家语料数据或用于大模型训练



资料来源：中信出版官网、华泰研究

随着大模型的应用，作为数据训练源头的版权 IP 将迎来较大需求，行业产能升级也将反哺版权 IP 价值。一方面，IP 版权内容为大模型数据训练源头，高质量数据库为大模型训练的刚需，IP 内容可帮助大模型公司实现高质量输入到输出的循环，随着大模型开发的进程，IP 版权或将迎来较大需求；另一方面，大模型将提升行业生产力效率，大幅缩短内容生产周期和降低成本，IP 有望突破原有的产能限制，不断累积价值效应，提升 IP 价值。据深交所互动易，以新文创 IP 内容的输出、运营和产业化布局为主业的百纳千成已成为百度文心一言首批生态合作伙伴，双方将在产品研发、标准制定等多个领域展开深化合作，引领内容生产和 IP 运营产业变革与升级。

视觉中国与文心一格在创作者赋能及版权保护方面展开合作。据深交所互动易，视觉中国通过与互联网平台的连接及赋能，在图文创作、视频生产、效果广告、电商广告、模版设计等业务场景提供图片、视频、音乐等优质正版内容服务。同时，视觉中国也与百度文心一格在创作者赋能和版权保护等方面展开多项合作，共探 AIGC 内容产业发展方向。

图表33：视觉中国图片库或为模型训练提供宝贵素材


资料来源：视觉中国官网、华泰研究

视频/虚拟人：缩短制作周期，提升交互能力

我们认为“如影”等生成式 AI 模型的落地将极大加速 AI 视频/虚拟人的发展，随着生成式 AI 与大语言模型的打通，视频/虚拟人制作周期将大大缩短，创作流程简化。同时，鉴于大模型在用户语言逻辑理解上的深化，虚拟人在场景应用中的识别感知能力和分析决策能力将显著提高，提升沟通时的交互能力，更准确地满足用户个性化需求。

“如影”打通大语言模型“商量”，打开新商业可能。据甲子光年，作为低门槛、快速打造数字人的 AI 数字人视频生成平台，如影未来能帮助电商直播企业、短视频营销企业、培训机构、娱乐企业等快速生成视频内容。我们认为，伴随着技术的迭代，AI 有望在视频的前期策划、中期拍摄、后期制作持续落地，降低视频生成的成本，提升效率，同时带来全新的创作思路，增强视频的创意性。

商汤发布会上，CEO 徐立展示了把“日日新 SenseNova”系列产品应用于直播间的可能，如用“商量 SenseChat”实时生成直播内容，用“如影 SenseAvatar”打造虚拟主播，用 3D 内容生成平台“琼宇 SenseSpace”和“格物 SenseThings”呈现线上场景和产品，打开了全新的商业可能。同时直播间能够 24 小时不停播，在大语言模型的加持下，虚拟人的互动性也极大提升。

图表34：“如影”可根据真人视频素材生成数字人分身


资料来源：商汤官网、华泰研究

图表35：“如影”可生成多种类视频

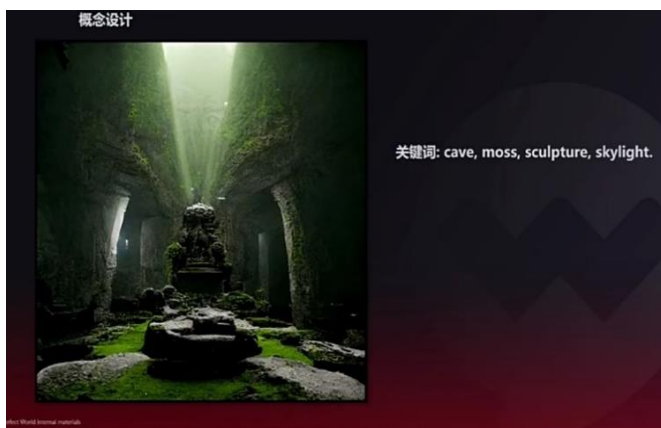

资料来源：商汤官网、华泰研究

游戏：赋能游戏素材生成，促进玩法创新

大模型可在 NPC 聊天内容、剧情大纲、场景生成等领域助力游戏生产，伴随多模态技术的发展，可能实现游戏生产的降本增效，促进玩法创新。

借助 AIGC 技术，游戏素材生成效率极大提升。完美世界技术美术主管 Gary Guo 在 2023 年 GDC 大会上分享利用 AIGC 技术进行美术设计，生成符合游戏场景的图标、图片、素材及 NPC 形象等。传统美术场景需要人工画师创作，但在 AIGC 技术辅助下场景设计、概念设计等能够在短时间内完成创作。目前主流应用方式还是通过 ChatGPT 叠加 MidJourney，通过 ChatGPT 可生成符合游戏开发者要求的 Prompt，再将 Prompt 输入 MidJourney 等 AI 绘图工具即可得到相关图片，可以通过不断调整 prompt 对图片进行进一步修改。

图表36：AI 生成场景概述



资料来源：GDC2023、华泰研究

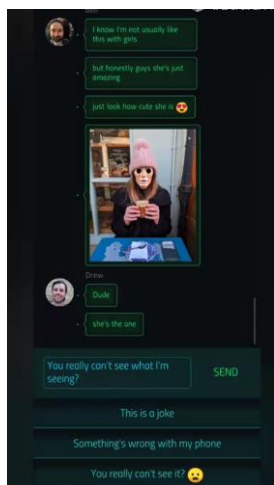
图表37：AI 生成角色概述



资料来源：GDC2023、华泰研究

大模型助力生成游戏剧情大纲等素材，提高创作效率，丰富剧情多样性。英国游戏工作室 ElectricNoir 旗下 Dark Mode 游戏主要采用 ChatGPT、DALL-E 2、AI 语音工具 Murf AI 三款 AI 工具制作，其中 ChatGPT 主要用于创建故事大纲和具体的人物对话，DALL-E 2 用于生成图像；中文在线海外产品 Chapters 和 My Escape 在进行接入 ChatGPT 测试，应用于故事创作生成、剧本生成及改编、交互聊天等方面。

图表38：Dark Mode 游戏第一章节试玩画面



资料来源：Gamelook、华泰研究

图表39：中文在线旗下虚拟恋爱互动游戏 My Escape



资料来源：App Store、华泰研究

大模型让游戏 NPC 产生更加多样化、符合设定的回答。我们认为，通过将游戏的世界观、剧情、人物对话等数据提供给大模型训练，游戏中的 NPC (non-player character) 便能够产生更加多样化、且符合设定的回答，极大程度提升游戏的可玩性、可拓展性。如汤姆猫已经尝试应用 GPT-3.5 模型进行 AI 语音互动产品功能原型测试，并已经将汤姆猫 IP 形象

植入产品功能原型当中，初步验证了相关技术实现的可行性；海外的沙盒游戏《Modbox》开发者通过 GPT-3 与 Repilca 模型实现 NPC 自由对话，不仅有语气还搭配丰富的肢体动作；日本开发商基于 ChatGPT 开发 AI 审讯游戏《ドキドキ AI 尋問ゲーム》中的 NPC 能够根据用户的输入进行实时生成回答，持续推动故事情节发展。

图表40：汤姆猫将虚拟形象接入类 ChatGPT 功能



资料来源：会说话的汤姆猫、华泰研究

图表41：审讯游戏《ドキドキ AI 尋問ゲーム》



资料来源：Gamelook、华泰研究

报告提及公司

图表42：报告提及公司表

公司名称	代码	公司名称	代码
阿里巴巴	BABA US	华为	未上市
OPPO	未上市	OpenAI	未上市
汤姆猫	300459 CH	中文在线	300364 CH
吉利汽车	175 HK	中信出版	300788 CH
ElectricNoir	未上市	视觉中国	000681 CH
奇瑞新智能	未上市	百度	BIDU US
豪墨自行	未上市	完美世界	002624 CH
可口可乐	KO US	Midjourney	未上市
波司登	3998 HK	V1	未上市
顾家家居	603816 CH	酷家乐	未上市
商汤科技	20 HK		

资料来源：Bloomberg、华泰研究

风险提示

1) AIGC 技术发展不及预期。AIGC（生成式 AI）是基于生成算法、训练数据、芯片算力，生成包括文本、音乐、图片、代码、视频等多样化的内容，目前仍然处于早期阶段，若技术的发展进度不及预期，则会进一步影响用户体验，降低沉浸感；

2) 政策监管风险。AI 发展迅速，相关法律法规、监管准则仍有较大的不确定性，若相关政策趋严，可能导致技术运用无法落地；

3) 商业化存在不确定性。AI 商业化落地受算力、算法、数据、场景、用户等多因素共同影响，需要大量资金投入。若商业化进度不及预期，可能进一步影响到研发资金的供给，导致技术发展缓慢，影响用户体验；

4) 本研报中涉及到未上市公司或未覆盖个股内容，均系对其客观公开信息的整理，并不代表本研究团队对该公司、该股票的推荐或覆盖。

免责声明

分析师声明

本人，朱璐，兹证明本报告所表达的观点准确地反映了分析师对标的证券或发行人的个人意见；彼以往、现在或未来并无就其研究报告所提供的具体建议或所表达的意见直接或间接收取任何报酬。

一般声明及披露

本报告由华泰证券股份有限公司（已具备中国证监会批准的证券投资咨询业务资格，以下简称“本公司”）制作。本报告所载资料是仅供接收人的严格保密资料。本报告仅供本公司及其客户和其关联机构使用。本公司不因接收人收到本报告而视其为客户。

本报告基于本公司认为可靠的、已公开的信息编制，但本公司及其关联机构（以下统称为“华泰”）对该等信息的准确性及完整性不作任何保证。

本报告所载的意见、评估及预测仅反映报告发布当日的观点和判断。在不同时期，华泰可能会发出与本报告所载意见、评估及预测不一致的研究报告。同时，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波动。以往表现并不能指引未来，未来回报并不能得到保证，并存在损失本金的可能。华泰不保证本报告所含信息保持在最新状态。华泰对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司不是 FINRA 的注册会员，其研究分析师亦没有注册为 FINRA 的研究分析师/不具有 FINRA 分析师的注册资格。

华泰力求报告内容客观、公正，但本报告所载的观点、结论和建议仅供参考，不构成购买或出售所述证券的要约或招揽。该等观点、建议并未考虑到个别投资者的具体投资目的、财务状况以及特定需求，在任何时候均不构成对客户私人投资建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。对依据或者使用本报告所造成的一切后果，华泰及作者均不承担任何法律责任。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。

除非另行说明，本报告中所引用的关于业绩的数据代表过往表现，过往的业绩表现不应作为日后回报的预示。华泰不承诺也不保证任何预示的回报会得以实现，分析中所做的预测可能是基于相应的假设，任何假设的变化可能会显著影响所预测的回报。

华泰及作者在自身所知情的范围内，与本报告所指的证券或投资标的不存在法律禁止的利害关系。在法律许可的情况下，华泰可能会持有报告中提到的公司所发行的证券头寸并进行交易，为该公司提供投资银行、财务顾问或者金融产品等相关服务或向该公司招揽业务。

华泰的销售人员、交易人员或其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。华泰没有将此意见及建议向报告所有接收者进行更新的义务。华泰的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。投资者应当考虑到华泰及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突。投资者请勿将本报告视为投资或其他决定的唯一信赖依据。有关该方面的具体披露请参照本报告尾部。

本报告并非意图发送、发布给在当地法律或监管规则下不允许向其发送、发布的机构或人员，也并非意图发送、发布给因可得到、使用本报告的行为而使华泰违反或受制于当地法律或监管规则的机构或人员。

本报告版权仅为本公司所有。未经本公司书面许可，任何机构或个人不得以翻版、复制、发表、引用或再次分发他人（无论整份或部分）等任何形式侵犯本公司版权。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并需在使用前获取独立的法律意见，以确定该引用、刊发符合当地适用法规的要求，同时注明出处为“华泰证券研究所”，且不得对本报告进行任何有悖原意的引用、删节和修改。本公司保留追究相关责任的权利。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。

中国香港

本报告由华泰证券股份有限公司制作，在香港由华泰金融控股（香港）有限公司向符合《证券及期货条例》及其附属法律规定的机构投资者和专业投资者的客户进行分发。华泰金融控股（香港）有限公司受香港证券及期货事务监察委员会监管，是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。在香港获得本报告的人员若有任何有关本报告的问题，请与华泰金融控股（香港）有限公司联系。

香港-重要监管披露

- 华泰金融控股（香港）有限公司的雇员或其关联人士没有担任本报告中提及的公司或发行人的高级人员。
- 有关重要的披露信息，请参华泰金融控股（香港）有限公司的网页 https://www.htsc.com.hk/stock_disclosure 其他信息请参见下方“美国-重要监管披露”。

美国

在美国本报告由华泰证券（美国）有限公司向符合美国监管规定的机构投资者进行发表与分发。华泰证券（美国）有限公司是美国注册经纪商和美国金融业监管局（FINRA）的注册会员。对于其在美国分发的研究报告，华泰证券（美国）有限公司根据《1934年证券交易法》（修订版）第15a-6条规定以及美国证券交易委员会人员解释，对本研究报告内容负责。华泰证券（美国）有限公司联营公司的分析师不具有美国金融监管（FINRA）分析师的注册资格，可能不属于华泰证券（美国）有限公司的关联人员，因此可能不受FINRA关于分析师与标的公司沟通、公开露面和所持交易证券的限制。华泰证券（美国）有限公司是华泰国际金融控股有限公司的全资子公司，后者为华泰证券股份有限公司的全资子公司。任何直接从华泰证券（美国）有限公司收到此报告并希望就本报告所述任何证券进行交易的人士，应通过华泰证券（美国）有限公司进行交易。

美国-重要监管披露

- 分析师朱珺本人及相关人士并不担任本报告所提及的标的证券或发行人的高级人员、董事或顾问。分析师及相关人士与本报告所提及的标的证券或发行人并无任何相关财务利益。本披露中所提及的“相关人士”包括FINRA定义下分析师的家庭成员。分析师根据华泰证券的整体收入和盈利能力获得薪酬，包括源自公司投资银行业务的收入。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或不时会以自身或代理形式向客户出售及购买华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）华泰证券研究所覆盖公司的证券/衍生工具，包括股票及债券（包括衍生品）。
- 华泰证券股份有限公司、其子公司和/或其联营公司，及/或其高级管理层、董事和雇员可能会持有本报告中所提到的任何证券（或任何相关投资）头寸，并可能不时进行增持或减持该证券（或投资）。因此，投资者应该意识到可能存在利益冲突。
- 本报告所载的观点、结论和建议仅供参考，不构成购买或出售所述证券的要约或招揽，亦不试图促进购买或销售该等证券。如任何投资者为美国公民、取得美国永久居留权的外国人、根据美国法律所设立的实体（包括外国实体在美国的分支机构）、任何位于美国的个人，该等投资者应当充分考虑自身特定状况，不以任何形式直接或间接地投资本报告涉及的投资者所在国相关适用的法律法规所限制的企业的公开交易的证券、其衍生证券及用于为该等证券提供投资机会的证券的任何交易。该等投资者对依据或者使用本报告内容所造成的一切后果，华泰证券股份有限公司、华泰金融控股（香港）有限公司、华泰证券（美国）有限公司及作者均不承担任何法律责任。

评级说明

投资评级基于分析师对报告发布日后6至12个月内行业或公司回报潜力（含此期间的股息回报）相对基准表现的预期（A股市场基准为沪深300指数，香港市场基准为恒生指数，美国市场基准为标普500指数），具体如下：

行业评级

- 增持：**预计行业股票指数超越基准
- 中性：**预计行业股票指数基本与基准持平
- 减持：**预计行业股票指数明显弱于基准

公司评级

- 买入：**预计股价超越基准15%以上
- 增持：**预计股价超越基准5%~15%
- 持有：**预计股价相对基准波动在-15%~5%之间
- 卖出：**预计股价弱于基准15%以上
- 暂停评级：**已暂停评级、目标价及预测，以遵守适用法规及/或公司政策
- 无评级：**股票不在常规研究覆盖范围内。投资者不应期待华泰提供该等证券及/或公司相关的持续或补充信息

法律实体披露

中国: 华泰证券股份有限公司具有中国证监会核准的“证券投资咨询”业务资格, 经营许可证编号为: 91320000704041011J
香港: 华泰金融控股(香港)有限公司具有香港证监会核准的“就证券提供意见”业务资格, 经营许可证编号为: AOK809
美国: 华泰证券(美国)有限公司为美国金融业监管局(FINRA)成员, 具有在美国开展经纪交易商业业务的资格, 经营业务许可编号为: CRD#:298809/SEC#:8-70231

华泰证券股份有限公司**南京**

南京市建邺区江东中路228号华泰证券广场1号楼/邮政编码: 210019

电话: 86 25 83389999/传真: 86 25 83387521

电子邮件: ht-rd@htsc.com

深圳

深圳市福田区益田路5999号基金大厦10楼/邮政编码: 518017

电话: 86 755 82493932/传真: 86 755 82492062

电子邮件: ht-rd@htsc.com

北京

北京市西城区太平桥大街丰盛胡同28号太平洋保险大厦A座18层/

邮政编码: 100032

电话: 86 10 63211166/传真: 86 10 63211275

电子邮件: ht-rd@htsc.com

上海

上海市浦东新区东方路18号保利广场E栋23楼/邮政编码: 200120

电话: 86 21 28972098/传真: 86 21 28972068

电子邮件: ht-rd@htsc.com

华泰金融控股(香港)有限公司

香港中环皇后大道中99号中环中心58楼5808-12室

电话: +852-3658-6000/传真: +852-2169-0770

电子邮件: research@htsc.com

<http://www.htsc.com.hk>**华泰证券(美国)有限公司**

美国纽约公园大道280号21楼东(纽约10017)

电话: +212-763-8160/传真: +917-725-9702

电子邮件: Huatai@htsc-us.com

<http://www.htsc-us.com>

©版权所有2023年华泰证券股份有限公司