



计算机行业研究

买入（维持评级）

行业研究

证券研究报告

计算机组

分析师：李可夫（执业 S1130525120009） 分析师：刘高畅（执业 S1130525120005）

likefu@gjzq.com.cn

liugaochang@gjzq.com.cn

GPT-6 ASTRA 发布，不只是 Coding

GPT-6 ASTRA 发布，多维度能力实现大幅升级。1) 计算机自主操作能力显著增强：模型可独立完成表单填写、CRM 更新、软件故障排查等大量真实电脑实操任务；OSWorld 2.0 模拟测试中单任务耗时相比前代降低约 47%，AI 不再局限输出文本代码，可直接操作软件跑完完整业务链路，向自主智能体迈出关键一步，释放企业办公自动化潜力。

2) 专业工作流、人机协作能力得到强化：模型能够遵循业务模板输出文档、PPT、PCB 方案、3D 场景工程文件等专业成品；支持任务中途叠加新增需求，最大程度保留原有任务约束，减少任务目标丢失，人机协同处理多步骤专业业务的实用性明显提升。3) 编程、数理科研、网络安全能力同步突破：编程评测得分明显提升，上线长任务上下文记忆实验功能，缓解超长任务信息压缩失真痛点；数学科研评测成绩大幅走高，能够参与测序校验、科研可视化等实操环节；网络安全能力达到关键阈值，正式版本针对高危能力实施安全管控，仅开放防御类安全相关工作。

第三方评测表现亮眼，多项基准实现跨越式提升。ARC Prize 评测显示，适配特定运行框架后 Astra 视觉交互推理得分接近满分，操作效率优于人类参与者中位数，但高分跑分不等同于实现 AGI，模型实际效果由模型本体与运行调度框架共同决定，基准测试结果不能完全等同于真实业务落地表现。产业端专业用户实测进一步佐证模型实力，在法律复杂文书处理、代码库理解、业务报告生成等真实业务场景展现实用价值，已经能够对商业产品形成实际赋能。

模型应用边界持续拓宽，商业化叙事有望跳出编码单一赛道。我们认为当前资本市场存在核心担忧，大模型商业化收入过度依赖编码场景，办公、法务、科研等非编码专业场景端到端交付能力不足，落地进度缓慢，或将压制行业长期天花板。GPT-6 Astra 迭代释放的核心信号，是非编码专业任务端到端交付能力实现实质性跃迁，推动模型从问答工具转变为任务执行主体，办公、法务、科研、硬件设计等非代码场景商业化进程有望加速。同时也要客观看待能力边界，法律、科研等高风险业务的模型输出仍需要专业人员审核校验，官方演示案例不等于行业规模化生产力。未来面向 B 端岗位的端到端任务交付能力，或将成为大模型商业化竞争的关键分水岭，国内厂商除编码能力外，也需要布局通用办公、行业工作流、计算机操作智能体等方向。

相关标的

大模型&云：智谱、MiniMax、阿里巴巴、腾讯控股、金山云、百度集团、优刻得、首都在线、网宿科技、云赛智联、青云科技等。

海外算力：东山精密、工业富联、胜宏科技、江海股份、中际旭创、中钨高新、蓝思科技、东阳光、火炬电子、三环集团、欧科亿、天孚通信、鼎泰高科、新易盛、领益智造、兆易创新、鹏鼎控股、唯科科技、天岳先进、大普微、源杰科技、麦格米特、景旺电子、英维克、京东方等；英特尔、SK 海力士、Lumentum、闪迪、铠侠、美光。

国内算力：寒武纪、海光信息、长鑫科技、东阳光、利通电子、中芯国际、华虹半导体、华达科技、摩尔线程、天数智芯、沐曦股份、壁仞科技、协创数据、禾盛新材、华丰科技、杰华特、京基智农、中科曙光、浪潮信息、利扬芯片、胜蓝股份、华勤技术、国科微、中国长城、晶科科技、罗曼股份、盈峰环境、芯原股份、晶科科技、亿田智能、豫能控股、星环科技、鸿日达、盛视科技、神州数码、润泽科技、大位科技、润建股份、奥飞数据、瑞晟智能、科华数据、潍柴重机、欧陆通、杰创智能、奥尼电子。

AI 应用：达梦数据、深信服、卓易信息、星环科技、广立微、绿盟科技、网宿科技、海天瑞声、合合信息、德才股份、美年健康、真爱美家、中控技术、迈富时、金蝶国际、康众医疗、用友网络、聚水潭、阜博集团、范式智能、汇量科技、金山办公、奔图科技、中国软件

风险提示

行业竞争加剧的风险；技术研发进度不及预期的风险；特定行业下游资本开支周期性波动的风险。



内容目录

一、GPT-6 ASTRA 发布，多维度能力实现大幅升级	3
1.1 模型分阶段灰度上线，多渠道逐步开放调用权限	3
1.2 计算机自主操作能力显著增强，任务耗时大幅缩减	3
1.3 专业工作流能力强化，模板遵循与多步骤协作能力提升	3
1.4 编程能力持续迭代，新增长任务上下文记忆实验性功能	3
1.5 数学与科研能力提升，网络安全能力同步增强	3
二、第三方评测表现亮眼，多项基准实现跨越式提升	4
2.1 ARC Prize 肯定 ASTRA 交互推理进展	4
2.2 来自产业专业用户的真实反馈，复杂业务场景实用性得到验证	5
三、模型应用边界持续拓宽，商业化叙事跳出编码单一赛道	5
3.2 市场核心担忧：商业化是否高度依赖编码赛道，非代码场景落地进度存疑	5
3.2 GPT-6 Astra 发布的核心信号：不只是编码变强，非编码专业场景能力迎来实质性跃迁	5
四、相关标的	5
五、风险提示	6

图表目录

图表 1： ARC-AGI-3 的视觉推理题	4
------------------------------	---



一、GPT-6 ASTRA 发布，多维度能力实现大幅升级

1.1 模型分阶段灰度上线，多渠道逐步开放调用权限

2026 年 9 月 4 日 OpenAI 正式发布 GPT-6 Astra，采取分阶段的灰度开放策略：优先定向向部分机构客户开放，后续数日逐步向 GPT Plus、Pro、Business、Enterprise 全量付费用户，同时开放 API、Microsoft Azure、AWS Bedrock 三大对外调用渠道。Pro、Business 以及 Enterprise 企业用户还可额外获取更高规格的 GPT-6 Astra Pro 版本。企业侧工作空间默认关闭 Astra 访问权限，需要管理员手动开启后才可使用，体现出对高能力模型企业部署的安全管控思路。。

1.2 计算机自主操作能力显著增强，任务耗时大幅缩减

计算机操作是本次 Astra 迭代的核心亮点，模型可独立完成大量真实电脑端业务流程：填写线上表单、更新 CRM 客户档案、整理日历日程、开展网络调研、邮件与文档摘要起草；同时覆盖科学数据分析、图表生成、网站搭建、前端质量校验、软件安装测试、屏幕故障排查等复杂实操任务。

在 OSWorld 2.0 延迟模拟测试中，GPT-6 Astra 得分 72.6%，单任务平均耗时约 40 分钟；前代 GPT-5.6 Sol 得分 65.7%，单任务平均耗时 75 分钟，单任务完成时间直接降低约 47%。AI 不再局限于输出文本、代码，而是可以直接操作软件完成完整业务链路，向自主智能体迈出关键一步，为企业端大量重复性办公流程自动化打开空间。

1.3 专业工作流程能力强化，模板遵循与多步骤协作能力提升

Astra 将电脑操作能力与面向专业环境的训练结合，用于多步骤工作流程，并生成文档、电子表格与演示文稿。发布文将遵循已有模板、匹配写作与视觉风格，以及筛选与任务相关的上下文，列为专业工作能力的具体改进方向。

发布页提供了几类可辨认的成果示例：在 KiCad 中进行印刷电路板布局，将电子原理图转化为可制造的 PCB；利用 OpenAI 演示文稿模板中的少量参考页，制作关于虚构模型 GPT-Gaia 的幻灯片；在 Blender 中建模住宅，并将其转化为 Unreal Engine 5 中的可漫游场景。上述内容均为官方展示案例。

在人机协作方面，官方称 Astra 能够利用上下文补全常规信息缺口，并在答案可能影响结果时提出针对性问题。在 Codex 中，模型可一边等待用户回复，一边继续不依赖该回复的工作；对于重要决定则等待用户输入。发布文还说明，Astra 能够将新增要求纳入原任务，减少因中途调整而丢失原有目标和约束的情况。

1.4 编程能力持续迭代，新增长任务上下文记忆实验性功能

编程依旧是 Astra 重点优化赛道，多项权威评测指标实现明显提升。Terminal-Bench 4.0 评测中 Astra 得分 57.9%，GPT-5.6 Sol 仅 37.3%；DeepSWE v1.1 评测 Astra 得分 74.1%，对比前代小幅提升。

针对长时间调试、大规模代码重构这类超长任务，OpenAI 为 Codex 配套全新实验功能：跨上下文窗口保存笔记，同时支持检索历史窗口全部信息。模型可以回溯过往消息、工具输出中的需求、测试结果，不完全依赖笔记压缩摘要保存信息。该功能现阶段处于可配置开启的实验阶段，官方计划数周后将其设置为 Astra 默认能力，缓解超长任务下上下文丢失、信息压缩失真的行业痛点。

1.5 数学与科研能力提升，网络安全能力同步增强

科研数学能力上，FrontierMath Tier4 (v2) 评测 Astra 得分 97.6%，GPT-5.6 Sol 为 83.0%；GPQA Diamond 评测 Astra 得分 96.0%，前代模型 94.6%。官方演示案例显示模型可直接调用专业科研软件完成测序质量校验、遗传变异可视化，真正介入科研实操环节，而不只是输出理论文本。

网络安全领域，Astra 达到 OpenAI Preparedness Framework 安全框架的 Critical 关键能力阈值。关闭生产防护的测试环境中，ExploitBench 评测拿到 100% 满分，GPT-5.6 Sol 仅 78.5%。面向普通用户的正式发布版本，支持安全代码审查、漏洞修补等防御性工作，但会拒绝生成漏洞利用 POC 等高风险任务，能力做分级管控，评测高分不等于向普通用户完全放开全部能力。



二、第三方评测表现亮眼，多项基准实现跨越式提升

2.1 ARC Prize 肯定 ASTRA 交互推理进展

ARC Prize 第三方评测结果显示，在 ARC-AGI-3 半私有测试集，标准运行框架下 Astra (max) 得分 62.7%；切换至支持跨请求推理状态保存、长对话压缩的 Provider Adapter 框架后，Astra (high) 得分高达 99.9%。

Provider Adapter 框架下，96.0% 的关卡 Astra 操作次数少于人类参与者中位数，平均每关操作次数降低 51.7%。ARC Prize 机构特别提示：基准测试接近满分，并不等同于实现 AGI，封闭测试环境无法完全复刻现实世界复杂多变的真实业务场景，跑分仅作为参考，不能直接等同于真实业务落地效果。不同运行框架带来巨大分数差异，也提示我们，模型实际产出效果，是模型本身能力 + 运行调度框架共同决定。

图表1：ARC-AGI-3 的视觉推理题



来源：ARC-AGI 官网，国金证券研究所



2.2 来自产业专业用户的真实反馈，复杂业务场景实用性得到验证

OpenAI 引用多家产业机构的实测反馈，印证 Astra 在真实专业任务中的表现。法律服务商 Harvey 应用研究负责人 Niko Grupen 反馈，Astra 处理复杂法律任务时，可以区分原始文件与既定业务记录，识别材料中缺少论据支撑的假设，并且把信息缺口转化为可落地的法律起草立场。

AI 开发公司 Cognition 研究高级副总裁 Silas Alberti 表示，电脑操作、文本写作、代码库理解能力的升级，直接优化 Devin 产品内部测试效果，输出视频可读性更强，业务报告更加简洁清晰。来自产业端的反馈，说明模型能力已经可以对下游产品形成实际赋能，而不只是停留在演示案例层面。

三、模型应用边界持续拓宽，商业化叙事跳出编码单一赛道

2026 年头部大模型厂商 ARR 持续快速增长，但我们认为当前市场存在大模型商业化的疑虑，ASTRA 有望消除部分担忧。

3.2 市场核心担忧：商业化是否高度依赖编码赛道，非代码场景落地进度存疑

我们认为，站在当前时间点，资本市场对全球大模型商业化存在一个普遍的核心担忧：现阶段大模型商业化收入很大一部分来自代码生成、代码辅助类 Coding 场景。编码需求付费意愿强、价值容易量化，是海内外模型厂商 ARR 增长的核心支柱。

但市场同时存在顾虑：除编码之外，办公、法律、科研、硬件设计、工业流程等大量非编码专业场景，当前模型大多仅能完成信息摘抄、初稿辅助，很难交付可以直接落地的完整成品，端到端闭环能力不足。市场担忧非编码领域商业化落地节奏会显著慢于编码赛道，大模型长期商业化天花板会被约束。若 B 端付费长期集中在软件开发赛道，整个大模型行业市场空间将受到压制，会影响市场对大模型企业长期价值的判断。

3.2 GPT-6 Astra 发布的核心信号：不只是编码变强，非编码专业场景能力迎来实质性跃迁

我们认为，GPT-6 Astra 的发布，为上述市场担忧提供了新的产业观测窗口。本次迭代中，编码能力固然继续提升，但更大的亮点恰恰集中在编码之外的各类专业工作交付能力：自主操作电脑完成表单填写、CRM 维护、日历管理；直接输出符合企业模板规范的 PPT、表格、正式文档；完成 PCB 硬件设计、3D 建模引擎联动、法律材料梳理、科研数据分析等大量非代码专业任务。这意味着，除软件研发之外，企业内部通用办公、法务、科研、硬件研发等岗位，都打开了 AI 降本增效以及商业付费的想象空间，大模型商业化并非只有 Coding 单一主线。

当然我们依旧需要客观看待能力边界：现阶段优先看好结果易核验、出错易修正的普通办公自动化场景；法律文书、科研结论、网络安全等高风险业务，模型输出必须经过专业人员审核，不能直接采信模型输出。官方演示案例不能直接等同于全行业规模化生产力提升。但产业趋势层面，Astra 证明非编码场景模型能力迎来拐点，办公、专业服务赛道商业化落地节奏有望加速，拓宽大模型整体商业化空间。

我们认为未来大模型竞争不局限于代码能力比拼，面向真实 B 端岗位的任务交付能力将成为拉开商业化差距的关键。国内厂商除打磨编码能力外，也需要在通用办公、多行业专业 workflow、计算机操作智能体方向持续投入，挖掘更广阔的商业化空间。

四、相关标的

AI 应用：达梦数据、深信服、卓易信息、星环科技、广立微、绿盟科技、网宿科技、海天瑞声、合合信息、德才股份、美年健康、真爱美家、中控技术、迈富时、金蝶国际、康众医疗、用友网络、聚水潭、阜博集团、范式智能、汇量科技、金山办公、奔图科技、中国软件

大模型&云：智谱、MiniMax、阿里巴巴、腾讯控股、金山云、百度集团、优刻得、首都在线、网宿科技、云赛智联、青云科技等。

海外算力：东山精密、工业富联、胜宏科技、江海股份、中际旭创、中钨高新、蓝思科技、东阳光、火炬电子、三环集团、欧科亿、天孚通信、鼎泰高科、新易盛、领益智造、兆易创新、鹏鼎控股、唯科科技、天岳先进、大普微、源杰科技、麦格米特、景旺电子、英维克、京东方等；英特尔、SK 海力士、Lumentum、闪迪、铠侠、美光。



国内算力：寒武纪、海光信息、长鑫科技、东阳光、利通电子、中芯国际、华虹半导体、华达科技、摩尔线程、天数智芯、沐曦股份、壁仞科技、协创数据、禾盛新材、华丰科技、杰华特、京基智农、中科曙光、浪潮信息、利扬芯片、胜蓝股份、华勤技术、国科微、中国长城、晶科科技、罗曼股份、盈峰环境、芯原股份、晶科科技、亿田智能、豫能控股、星环科技、鸿日达、盛视科技、神州数码、润泽科技、大位科技、润建股份、奥飞数据、瑞晟智能、科华数据、潍柴重机、欧陆通、杰创智能、奥尼电子。

五、风险提示

行业竞争加剧的风险：模型厂商若竞争加剧可能影响利润率。

技术研发进度不及预期的风险：模型能力迭代若不及预期可能影响大模型厂商商业化。

特定行业下游资本开支周期性波动的风险：云厂商、运营商等客户对 AI 服务器与算力基础设施投资力度波动风险。



行业投资评级的说明：

买入：预期未来 3—6 个月内该行业上涨幅度超过大盘在 15%以上；

增持：预期未来 3—6 个月内该行业上涨幅度超过大盘在 5%—15%；

中性：预期未来 3—6 个月内该行业变动幅度相对大盘在 -5%—5%；

减持：预期未来 3—6 个月内该行业下跌幅度超过大盘在 5%以上。



特别声明：

国金证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。

本报告版权归“国金证券股份有限公司”（以下简称“国金证券”）所有，未经事先书面授权，任何机构和个人均不得以任何方式对本报告的任何部分制作任何形式的复制、转发、转载、引用、修改、仿制、刊发，或以任何侵犯本公司版权的其他方式使用。经过书面授权的引用、刊发，需注明出处为“国金证券股份有限公司”，且不得对本报告进行任何有悖原意的删节和修改。

本报告的产生基于国金证券及其研究人员认为可信的公开资料或实地调研资料，但国金证券及其研究人员对这些信息的准确性和完整性不作任何保证。本报告反映撰写研究人员的不同设想、见解及分析方法，故本报告所载观点可能与其他类似研究报告的观点及市场实际情况不一致，国金证券不对使用本报告所包含的材料产生的任何直接或间接损失或与此有关的其他任何损失承担任何责任。且本报告中的资料、意见、预测均反映报告初次公开发布时的判断，在不作事先通知的情况下，可能会随时调整，亦可因使用不同假设和标准、采用不同观点和分析方法而与国金证券其它业务部门、单位或附属机构在制作类似的其他材料时所给出的意见不同或者相反。

本报告仅为参考之用，在任何地区均不应被视为买卖任何证券、金融工具的要约或要约邀请。本报告提及的任何证券或金融工具均可能含有重大的风险，可能不易变卖以及不适合所有投资者。本报告所提及的证券或金融工具的价格、价值及收益可能会受汇率影响而波动。过往的业绩并不能代表未来的表现。

客户应当考虑到国金证券存在可能影响本报告客观性的利益冲突，而不应视本报告为作出投资决策的唯一因素。证券研究报告是用于服务具备专业知识的投资者和投资顾问的专业产品，使用时必须经专业人士进行解读。国金证券建议获取报告人员应考虑本报告的任何意见或建议是否符合其特定状况，以及（若有必要）咨询独立投资顾问。报告本身、报告中的信息或所表达意见也不构成投资、法律、会计或税务的最终操作建议，国金证券不就报告中的内容对最终操作建议做出任何担保，在任何时候均不构成对任何人的个人推荐。

在法律允许的情况下，国金证券的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，并可能为这些公司正在提供或争取提供多种金融服务。

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、发布该研究报告的人员。国金证券并不因收件人收到本报告而视其为国金证券的客户。本报告对于收件人而言属高度机密，只有符合条件的收件人才能使用。根据《证券期货投资者适当性管理办法》，本报告仅供国金证券股份有限公司客户中风险评级高于 C3 级（含 C3 级）的投资者使用；本报告所包含的观点及建议并未考虑个别客户的特殊状况、目标或需要，不应被视为对特定客户关于特定证券或金融工具的建议或策略。对于本报告中提及的任何证券或金融工具，本报告的收件人须保持自身的独立判断。使用国金证券研究报告进行投资，遭受任何损失，国金证券不承担相关法律责任。

若国金证券以外的任何机构或个人发送本报告，则由该机构或个人为此发送行为承担全部责任。本报告不构成国金证券向发送本报告机构或个人的收件人提供投资建议，国金证券不为此承担任何责任。

此报告仅限于中国境内使用。国金证券版权所有，保留一切权利。

上海

电话：021-80234211

邮箱：researchsh@gjzq.com.cn

邮编：201204

地址：上海浦东新区芳甸路 1088 号

紫竹国际大厦 5 楼

北京

电话：010-85950438

邮箱：researchbj@gjzq.com.cn

邮编：100005

地址：北京市东城区建国门内大街 26 号

新闻大厦 8 层南侧

深圳

电话：0755-86695353

邮箱：researchsz@gjzq.com.cn

邮编：518000

地址：深圳市福田区金田路 2028 号皇岗商务中心

18 楼 1806



【小程序】
国金证券研究服务



【公众号】
国金证券研究