

互联网&软件&半导体

北京调研总结：算力端需求旺盛，AI 应用迅速起量

我们于 9 月 14-18 日在北京开展了 TMT 系列公司调研，涵盖大模型、互联网、AI 基础设施及半导体硬件的相关公司，本次调研的主要收获包括：1) 算力的获取能力与成本是国内 AI 公司的共同约束，各公司在持续扩充算力供给、优化基础设施效率来进一步打开变现空间；2) 对于传统互联网公司而言（非云和大模型公司），AI 相关的投入短期仍主要以费用的形式体现，投入已在员工效率提升及生态体系效率提升中看到成效，但可量化的 ROI 产出或仍待时日；3) 半导体板块方面，继续看好利基型存储和设备产业链，AI 需求外溢与海外大厂供给收缩形成共振。在 AI 的使用场景中，我们认为中国在编程领域仍有较大渗透空间，网络安全场景起量迅速，未来金融、法律、工业制造、生命科学均有望成为新的行业拓展方向。互联网及软件行业建议持续关注基础模型中已证实技术能力且变现场景持续拓宽的大模型公司、龙头的多模态模型厂商、及 AI 应用持续深化赋能业务运营效率提升的互联网公司；半导体继续看好利基型存储和设备产业链的投资价值。

- **基础模型：扩容算力供给，拓宽变现场景。** 我们整体观察到基础模型厂商及云服务厂商迅速扩容算力供给以满足增量需求的趋势，Coding 场景快速起量、Cowork 持续渗透支撑 ARR 增长。而在行业拓展方面，除 Coding 和 Cowork 的场景外，垂类场景中，法律场景目前已经基本验证，未来变现有望深化；金融行业正在渗透，后续可关注的行业包括教育、电商和科研。在利润率层面，模型自身的优化（包括数据、算法和工程优化）、基础设施效率提升、及模型合理提价有望推动大模型厂商毛利率优化，而营收快速起量则有望支撑经营杠杆释放和未来亏损收窄。
- **视频模型：市场空间广阔。** 我们调研视频模型行业的核心收获包括：1) AI 视频模型行业仍然处于早期发展阶段，行业参与者众多，但充足的市场空间支撑行业头部厂商持续快速增长；2) 分场景来看，AI 短剧在过去两年发展迅速并已经达到较高渗透率，但是视频模型在广告、专业影视、游戏等场景或仍有较大的渗透率提升空间；3) 视频模型厂商在不同场景建立了差异化竞争优势，字节旗下 Seedance 凭借领先的视频生成效果和创作者生态在短剧场景具备一定优势，快手可灵和 MiniMax H3 则凭借视频生成的稳定性与可控性（例如稳定准确的文本字幕生成）在广告场景具备一定优势；4) 得益于模型架构优化与规模效应，头部视频模型厂商毛利率或仍有较大提升空间。
- **半导体：继续看好利基型存储和设备产业链。** AI 需求外溢与海外大厂供给收缩形成共振，利基存储“量、价、利”同步向上且景气周期显著拉长，国产替代窗口同步打开，具体体现为：1) 需求由 AI 单点扩散为全链条紧缺；2) 涨价幅度与向下游传导的能力均超预期；3) 供给刚性强化了价格刚性；4) 订单与稼动率同步走强，设备端最先感知；5) 国产替代加速导入；6) 竞争格局因紧缺而阶段性缓和。投资建议方面，建议沿存储与设备两条主线积极布局。存储方面，推荐关注“价格维持高位 + 产能与代工锁定能力”，中长期看国际大厂退出带来的份额与格局优化构成核心成长逻辑；设备方面，订单高增与交付受限并存，短期亮点在上游零部件国产化（业绩兑现早于整机），中期看随着明年洁净室陆续建成释放，景气度有望维持高位。

优于大市
(维持)

中国 TMT 行业

贺赛一, CFA
(852) 3916 1739
hesaiyi@cmbi.com.hk

周志敏
(852) 3850 5268
zhouzhimin@cmbi.com.hk

陶冶, CFA
(852) 3850 5226
franktao@cmbi.com.hk

陆文韬, CFA
luwentao@cmbi.com.hk

张元圣
(852) 3761 8727
kevinzhang@cmbi.com.hk

郭远宁
(852) 3916 3715
aaronguo@cmbi.com.hk

郭书音
(852) 3916 3716
guoshuyin@cmbi.com.hk

基础大模型：算力扩容支撑商业化增长，海外合作拓展变现渠道

基础大模型公司中我们主要走访了智谱华章（2513 HK，买入）。公司于 9 月 16 日召开公开电话会议将 2026 年年底的 ARR 目标提升至 30 亿美元（前值：1H26 业绩会指引 24 亿美元），主要支撑因素包括：1) Coding 场景起量迅速且尚未见到瓶颈，且算力紧缺的情况在融资后边际缓和；2) co-work 整体进展快于预期，目前行业订单金额已经超过 10 亿元，尽管确收或需要在 1-2 年的时间维度逐步进行；3) 云服务商合作拓展，公司已与多家海外大型云服务商及国内头部云服务商签署收入分成协议，相关收入预计从 10 月开始确认。

短期内算力供给已不再是收入增长的主要约束。公司 7 月再融资所得资金大部分用于算力扩容，9 月再融资及可转债安排合计 50 亿美元也主要投向算力资源。目前公司算力几乎全部采用租赁模式，训练与推理资源大致按四成、六成配置，并通过混合算力池动态调度提升利用效率。GLM-5.3 Flash 已在 10 万级国产卡集群跑通，公司未来将持续推进国产芯片适配、异构优化及自主运维集群探索。公司认为，在 7 月和 9 月两轮融资过后，短期内算力供给已经不再是公司收入增长的主要约束，且大规模算力集群将支持智谱下一代大基座模型的预训练及 RSI（Recursive Self-Improvement，递归式自我改进）研究。

模型方面，公司下一代产品预计于 4Q26 发布。公司强调参数与数据规模的合理匹配，同时布局 Flash 和高性能模型，兼顾不同用户的智能、速度及价格需求，并通过 Z Code、第三方 Coding Agent 合作及行业客户积累优质任务数据。

应用端，在行业和场景拓展层面，智谱正由 Coding 向 Cowork 及垂直行业延伸，API 仍是核心的商业化模式。Cowork 通过嵌入 ERP、CRM 等系统，打通跨部门流程，重点服务规则明确、流程标准化的工作。网络安全垂类应用层面，目前需求来自漏洞发现和代码审查，部分企业已有上亿元安全预算，但仍缺合适产品，公司致力于更好的满足用户的需求。法律场景目前已经基本验证，未来变现有望深化；金融行业正在渗透，后续可关注的行业包括教育、电商和科研。

盈利端的改善方面，公司提及改善将主要依靠模型自身的优化（包括数据、算法和工程三方面进行成本优化）、硬件效率提升、及模型合理提价来达成毛利率的优化，同时营收快速起量有望支撑经营杠杆提升和未来亏损收窄。

视频模型：市场空间广阔

我们调研了视频模型行业相关公司，核心要点包括：1) AI 视频模型行业仍然处于早期发展阶段，行业参与者众多，但充足的市场空间支撑行业头部厂商持续快速增长；2) 分场景来看，AI 短剧在过去两年发展迅速并已经达到较高渗透率，但是视频模型在广告、专业影视、游戏等场景仍然有较大的渗透率提升空间；3) 视频模型厂商在不同场景建立了差异化竞争优势，字节旗下 Seedance 凭借领先的视频生成效果和创作者生态在短剧场景具备一定优势，快手可灵和 MiniMax H3 则凭借视频生成的稳定性与可控性（例如稳定准确的文本字幕生成）在广告场景具备一定优势；4) 得益于模型架构优化与规模效应，我们认为头部视频模型厂商毛利率（如可灵毛利率当前约 20%）仍有较大提升空间。

全球 AI 视频模型市场空间广阔，快手（1024 HK，买入）测算全球视频内容创作市场空间超 1400 亿美元，其中 AI 视频模型渗透率仍然处于较低水平。尽管当前行业玩家众多，但得益于视频模型渗透率快速提升及市场快速增长，头部模型厂商仍然维持快速增长。可灵 AI 2Q26 收入达到 8.5 亿元，同比增长超过 200%，9-10 月可灵或迎来重大版本更新，模型能力对标 Seedance 2.5 的同时保持价格优势，有望进一步推动可灵 ARR 快速增长。

图 1: 全球文生视频模型榜单

Rank	Range	Creator	Model	Elo	95% CI	Samples	Released	API Pricing ¹
1	1-3	 Google	Gemini Omni Flash	1233	-7/7	15,172	May 2026	\$6.00 /min
2	1-3	 Alibaba	Wan 3.0	1229	-9/9	6,011	Aug 2026	\$12.00 /min
3	1-4	 Fal	Minimax H3 Max (post-trained by fal)	1227	-9/9	5,689	Aug 2026	\$2.40 /min
4	3-4	 MiniMax	MiniMax H3 	1220	-8/8	8,602	Jul 2026	\$780 /min
5	5	 ByteDance Seed	Dreamina Seedance 2.0 720p	1210	-6/6	20,345	Mar 2026	\$9.07 /min
6	6-7	 Sand.ai	MAGI-2 Preview (0912)	1156	-9/9	5,103	Sep 2026	Coming soon
7	7-8	 Alibaba	Wan2.7-260612	1149	-6/6	16,694	Jun 2026	\$9.00 /min
8	7-8	 Alibaba-ATH	HappyHorse-1.1	1147	-6/6	16,670	Jun 2026	\$9.90 /min
9	9	 Alibaba-ATH	HappyHorse-1.0	1119	-7/7	7,618	Apr 2026	\$13.20 /min
10	10-12	 Alibaba	Wan 2.7	1098	-7/7	4,856	Apr 2026	\$9.00 /min
11	10-14	 Skywork AI	SkyReels V4	1095	-7/7	4,991	Mar 2026	\$21.00 /min
12	10-13	 KlingAI	Kling 3.0 1080p (Pro)	1095	-6/6	18,305	Feb 2026	\$20.16 /min

资料来源: Artificial Analysis, 招银国际环球市场

短剧场景下, 据广电总局网络视听节目管理司, 2026 年前 8 个月中国上线微短剧 43 万部, 达到去年全年的 13 倍, 其中 AI 剧集占比超过 90%。从供给侧看, 视频模型在短剧制作中的渗透率已经达到较高水平。从竞争维度看, 字节旗下 Seedance 凭借领先的视频生成效果和创作者生态在短剧场景具备一定优势。

广告及中长影视场景下, 企业或创作者通常对视频模型的画面质量、视频生成的稳定性与一致性提出更高要求, 快手可灵与 MiniMax H3 呈现一定优势。同时未来可灵也会加强与专业创作者和版权方的合作, 加速视频模型在专业影视创作领域的渗透率提升。

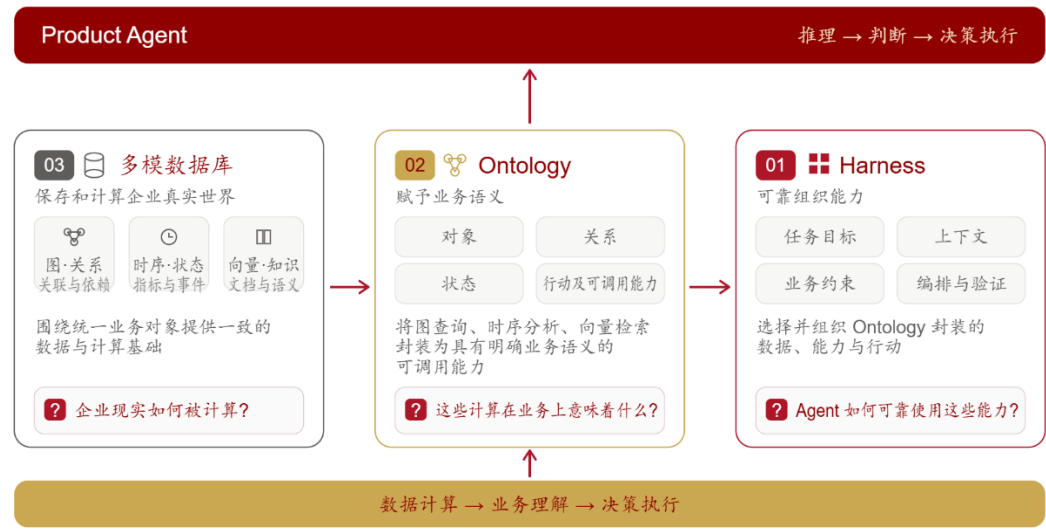
企业 AI 智能体: 加速落地

随着模型能力的提升与企业 AI Harness 框架的完善, 国内企业 AI 智能体需求快速增长。海致科技 (2706 HK, 买入) 目前已形成完善的产业级 AI 智能体框架: 1) 底层多模数据库为企业 AI 智能体提供数据与计算基础; 2) 中层本体 (Ontology) 体系将图查询、时序分析、向量检索等功能封装为有明确业务语义的可调用能力; 3) 上层 Harness 框架为企业智能体提供任务目标、业务约束、上下文与编排能力。

此外, 海致科技持续加强 FDE (Forward Deployed Engineer) 能力以加速产业智能体落地。区别于传统软件定制化开发, 海致 FDE: 1) 帮助企业客户发现并定义企业智能体的开发需求, 通过与企业的业务团队交流, 将企业内专家知识经验抽象并封装为企业智能体的能力; 2) 通过解决标杆客户的真实场景需求, 海致抽象出的企业智能体能力可复用于同行业其他客户, 后续同行业客户项目的交付效率及毛利率有望持续提升。

海致科技 1H26 总营收同比增长 70%至 2.96 亿元 (FY25: 同比增长 23%), 增长提速主要得益于 Atlas AI 智能体收入强劲增长 (同比增长 136%)。我们看好公司持续受益于企业客户对 AI 智能体的强劲需求, 主要得益于: 1) 行业领先的多模数据库能力, 而多模数据库有望成为企业 AI 智能体时代的核心数据基础设施; 2) 本体与 Atlas Harness 框架的持续完善与平台化能力; 3) 广泛的行业与用户覆盖。

图 2: 海致科技: 产业 AI 智能体解决方案



资料来源: 公司资料, 招银国际环球市场

中科闻歌 (1956 HK, 未评级) 则围绕 DOMA 统一架构形成从数据、本体、模型到智能体及决策的完整产品体系, 包括海汇 TokSea、DIP 决策智能平台、雅意 YAYI、ScienceOne、决策机 Decitron 及龙工 Clawworks, 重点提升模块复用、标准化部署和持续运营能力。此外, 公司持续加强 AI for Science 布局, 公司 9 月 17 日正式发布 SciencePro 磐石智研决策平台, 将决策 AI 能力延伸至科研场景。SciencePro 以中国科学院统一部署研发的磐石科学基础大模型为核心技术底座, 将模型、数据、工具、Skills 和科研智能体组织在统一环境中, 推出创新评价、ScienceHub 和科研智能体广场等核心能力, 为科研决策提供一体化支撑。

公司 1H26 新签合同金额达到 3.55 亿元; 报告期内服务客户约 290 家。得益于 AI 提升项目交付效率, 公司平均交付周期由 2023 年的 185 天缩短至 72 天, 毛利率改善的同时亏损显著收窄。

图 3: 中科闻歌: 全栈 AI 产品体系布局



资料来源: 公司资料, 招银国际环球市场

互联网公司 AI 应用及业务进展更新

■ 京东 (JD US, 买入)

京东在 AI 的使用方面主要包括：1) C 端：智能导购；2) B 端：助力采销提效、提供更多的商家工具及助力内部日常工作流提效。在 C 端的智能导购层面，新版京东 App 整体 AI 化，核心是把智能导购嵌入购前、购中、购后每个环节，目标是意图识别与转化率提升。在 B 端的用例方面，此前京东采销大量时间被商品上架、商详、商品描述占用，这些工作正在或已被 AI 替代，释放出的时间助力采销团队转向与品牌商建立关系、改善采购成本、发现趋势性品类。

此外，在核心业务的增长展望方面，公司指引京东零售业务长期高个位数经营利润率目标不变，驱动来自品类结构优化、三方平台业务占比提升、品类毛利率优化以及营销与履约提效，短期拖累则是 AI 带来的研发费用上涨。在海外拓展方面，公司将持续高效推进 Joybuy 业务的海外扩张以打开长期营收和盈利的增长前景。

■ 贝壳 (BEKE US, 买入)

贝壳在 AI 方面的探索主要集中于“服务流程重构”和“优秀经纪人能力复制”。公司致力于优化和重构 ACN (Agent Cooperation Network, 经纪人合作网络) 体系，以进一步推动组织效率的提升与工作流程的优化，适应 AI 时代发展的新需求。公司将借力 AI，通过“AI+人”服务管家帮助客户澄清需求以提升商机质量；并自研 AI harness 平台，结合公司自有数据、经验和知识搭建，目标是把优秀经纪人脑中的能力提炼、复制到更多人身上，并把地域化的服务能力泛化。此外，在 AI 的支持下，公司还将探索个性化的考核与决策，按区域及客户特点设置差异化考核以优化和提升管理运营效率。

我们认为此前 2Q26 的业绩已经清晰地证明了贝壳盈利能力正越来越多由单店及单经纪人产出提升，而非网络规模扩张所驱动。与收入规模大致相当的 2Q21 相比，2Q26 Non-GAAP 净利润高出 95%，店均交易单量提升 14%，人均房产交易单量提升 50%。我们认为，一个稳定但持续优化的服务网络、更高的门店及经纪人生产力，以及新业务更加健康的单位经济模型，有望在房地产市场复苏仍不均衡的环境下持续支撑公司的盈利能力。

半导体：继续看好利基型存储和设备产业链

本次调研的核心结论为 AI 需求外溢与海外大厂供给收缩形成共振，利基存储“量、价、利”同步向上且景气周期显著拉长，国产替代窗口同步打开，具体体现为：1) 需求由 AI 单点扩散为全链条紧缺。海外大厂将产能与设备资源集中投向 HBM、DDR5 等云端高端产品，主动退出利基存储（利基 DRAM、SLC NAND、NOR），腾出的市场空间被国内厂商承接。大厂退出利基市场属长期战略，周期结束后也不会重返低毛利市场，打开了利基型存储市场的成长空间。2) 涨价幅度与向下游传导的能力均超预期。DDR4、SLC NAND、SRAM 等利基型存储依然保持量价齐升的势头。目前看下游传导顺畅：高端手机、AI 服务器、高价扫地机尚有承接能力。3) 供给刚性强化了价格刚性。成熟存储产线扩产周期较长，且设备资源被 AI 主流需求挤占，利基存储 2026 年基本没有新增产能，预计新产能最早 2027H2–2028 年释放，2027 年之前利基产能持续偏紧。4) 订单与稼动率同步走强，设备端最先感知。受 AI 需求及其外溢效应拉动，存储、逻辑、成熟制程需求均保持旺盛。晶圆厂稼动率普遍高企，全球 WFE 市场规模预测年内多次上调，国产替代需求依然迫切。业内设备交期普遍拉长，在手订单饱满。5) 国产替代加速导入。全球供需紧张及地缘政治影响的双重驱动下，设备及存储的国产化替代都在加速演进，建议把握窗口期积极布局。6) 竞

争格局因紧缺而阶段性缓和。供需紧张背景下，全行业产能不足，短期内难现激烈竞争，毛利率有望维持高位稳定的状态。

投资建议：建议沿存储与设备两条主线积极布局。存储方面，涨价周期被拉长但斜率放缓，应关注“价格维持高位 + 产能与代工锁定能力”；中长期看，国际大厂退出带来的份额与格局优化构成核心成长逻辑。设备方面，订单高增与交付受限并存，短期亮点在上游零部件国产化（业绩兑现早于整机），中期看随着明年洁净室陆续建成释放，景气度有望维持高位。

免责声明及披露

分析员声明

负责撰写本报告的全部或部分内容的分析员，就本报告所提及的证券及其发行人做出以下声明：（1）发表于本报告的观点准确地反映有关他们个人对所提及的证券及其发行人的观点；（2）他们的薪酬在过往、现在和将来与发表在报告上的观点并无直接或间接关系。

此外，分析员确认，无论是他们本人还是他们的关联人士（按香港证券及期货事务监察委员会操作守则的相关定义）（1）并没有在发表研究报告 30 日前处置或买卖该等证券；（2）不会在发表报告 3 个工作日内处置或买卖本报告中提及的该等证券；（3）没有在有关香港上市公司内任职高级人员；（4）并没有持有有关证券的任何权益。

招银国际环球市场或其关联机构曾在过去 12 个月内与本报内所提及发行人有投资银行业务的关系。

招银国际环球市场投资评级

买入	：股价于未来 12 个月的潜在涨幅超过 15%
持有	：股价于未来 12 个月的潜在变幅在-10%至+15%之间
卖出	：股价于未来 12 个月的潜在跌幅超过 10%
未评级	：招银国际证券并未给予投资评级

招银国际环球市场行业投资评级

优于大市	：行业股价于未来12 个月预期表现跑赢大市指标
同步大市	：行业股价于未来12 个月预期表现与大市指标相若
落后大市	：行业股价于未来 12 个月预期表现跑输大市指标

招银国际环球市场有限公司

地址: 香港中环花园道 3 号冠君大厦 45 楼

电话: (852) 3900 0888

传真: (852) 3900 0800

招银国际环球市场有限公司(“招银国际环球市场”)为招银国际金融有限公司之全资附属公司(招银国际金融有限公司为招商银行之全资附属公司)

重要披露

本报告内所提及的任何投资都可能涉及相当大的风险。报告所载数据可能不适合所有投资者。招银国际环球市场不提供任何针对个人的投资建议。本报告没有把任何人的投资目标、财务状况和特殊需求考虑进去。而过去的表现亦不代表未来的表现，实际情况可能和报告中所载的大不相同。本报告中所提及的投资价值或回报存在不确定性及难以保证，并可能会受目标资产表现以及其他市场因素影响。招银国际环球市场建议投资者应该独立评估投资和策略，并鼓励投资者咨询专业财务顾问以便作出投资决策。

本报告包含的任何信息由招银国际环球市场编写，仅为本公司及其关联机构的特定客户和其他专业人士提供的参考数据。报告中的信息或所表达的意见皆不可作为或被视为证券出售要约或证券买卖的邀请，亦不构成任何投资、法律、会计或税务方面的最终操作建议，本公司及其雇员不就报告中的内容对最终操作建议作出任何担保。我们不对因依赖本报告所载资料采取任何行动而引致之任何直接或间接的错误、疏忽、违约、不谨慎或各类损失或损害承担任何的法律责任。任何使用本报告信息所作的投资决策完全由投资者自己承担风险。

本报告基于我们认为可靠且已经公开的信息，我们力求但不担保这些信息的准确性、有效性和完整性。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，可能会随时调整，且不承诺作出任何相关变更的通知。本公司可发布其它与本报告所载资料及/或结论不一致的报告。这些报告均反映报告编写时不同的假设、观点及分析方法。客户应该小心注意本报告中所提及的前瞻性预测和实际情况可能有显著区别，唯我们已合理、谨慎地确保预测所用的假设基础是公平、合理。招银国际环球市场可能采取与报告中建议及/或观点不一致的立场或投资决定。

本公司或其附属关联机构可能持有报告中提到的公司所发行的证券头寸并不时自行及/或代表其客户进行交易或持有该等证券的权益，还可能与这些公司具有其他投资银行相关业务联系。因此，投资者应注意本报告可能存在的客观性及利益冲突的情况，本公司将不会承担任何责任。本报告版权仅为本公司所有，任何机构或个人于未经本公司书面授权的情况下，不得以任何形式翻版、复制、转售、转发及或向特定读者以外的人士传阅，否则有可能触犯相关证券法规。如需索取更多有关证券的信息，请与我们联系。

对于接收此份报告的英国投资者

本报告仅提供给符合(I) 不时修订之英国 2000 年金融服务及市场法令 2005 年(金融推广)令(“金融服务令”)第 19(5) 条之人士及(II) 属金融服务令第 49(2) (a) 至(d) 条(高净值公司或非公司社团等)之机构人士，未经招银国际环球市场书面授权不得提供给其他任何人。

对于接收此份报告的美国投资者

招银国际环球市场不是在美国的注册经纪交易商。因此，招银国际环球市场不受美国就有关研究报告准备和研究分析员独立性的规则的约束。负责撰写本报告的全部或部分内容的分析员，未在美国金融业监管局(“FINRA”)注册或获得研究分析师的资格。分析员不受旨在确保分析师不受可能影响研究报告可靠性的潜在利益冲突的相关 FINRA 规则的限制。本报告仅提供给美国 1934 年证券交易法(经修订) 规则 15a-6 定义的“主要机构投资者”，不得提供给其他任何人。接收本报告之行为即表明同意接受协议不得将本报告分发或提供给任何其他人。接收本报告的美国收件人如想根据本报告中提供的信息进行任何买卖证券交易，都应仅通过美国注册的经纪交易商来进行交易。

对于在新加坡的收件人

本报告由 CMBI (Singapore) Pte. Limited (CMBISG) (公司注册号 201731928D) 在新加坡分发。CMBISG 是在《财务顾问法案》(新加坡法例第 110 章)下所界定，并由新加坡金融管理局监管的豁免财务顾问公司。CMBISG 可根据《财务顾问条例》第 32C 条下的安排分发其各自的外国实体，附属机构或其他外国研究机构编制的报告。如果报告在新加坡分发给非《证券与期货法案》(新加坡法例第 289 章)所定义的认可投资者，专家投资者或机构投资者，则 CMBISG 仅会在法律要求的范围内对这些人士就报告内容承担法律责任。新加坡的收件人应致电 (+65 6350 4400) 联系 CMBISG，以了解由本报告引起或与之相关的事宜。